INTRODUCCIÓN AL
ÁLGEBRA LINEAL

Serge Lang



ADDISON-WESLEY IBEROAMERICANA

INTRODUCCIÓN AL
ÁLGEBRA LINEAL

Serge Lang

Versión en español de

Miguel Lara Aparicio
Universidad Nacional Autónoma de México

Con la colaboración de

Guillermo Hansen
Universidad de Buenos Aires, Argentina



Addison-Wesley Iberoamericana

Argentina • Brasil • Chile • Colombia • Ecuador • España
Estados Unidos • México • Perú • Puerto Rico • Venezuela

Versión en español de la obra titulada *Introduction to Linear Algebra, Second Edition*, de Serge Lang, publicada originalmente en inglés por Springer-Verlag New York Inc., ©1970, 1986 por Springer-Verlag New York Inc. La primera edición en inglés fue publicada en 1970 por Addison-Wesley Publishing Company Inc.

Obra compuesta y formada mediante el sistema TEX por el Taller Lima, México.

Impreso en Estados Unidos. *Printed in U.S.A.*

ISBN 0-201-62912-7
6 7 8 9 10-CRS-99 98 97 96

Prefacio

Este libro se ideó como un texto breve para un curso semestral de álgebra lineal. Excepto por algún ejemplo o ejercicio ocasional, la lógica del texto es independiente del cálculo, y se podría utilizar desde los primeros cursos. En la práctica, mi intención es que lo utilicen principalmente los estudiantes que hayan cursado dos o tres semestres de cálculo. También se podría impartir el curso en forma simultánea con el primer curso de cálculo, o inmediatamente después de él.

He incluido algunos ejemplos de espacios vectoriales de funciones, aunque quienes deseen concentrarse en forma exclusiva en el estudio del espacio euclidiano podrían omitirlos del todo sin afectar la compresión del resto del libro. Además, si a algún lector le desagrada trabajar con $n = n$, siempre puede suponer que $n = 1, 2$ ó 3, y eliminar otras interpretaciones. Sin embargo, dicho lector debería observar que el uso de $n = n$ simplifica algunas fórmulas, abreviándolas por ejemplo, y debería tratar de alcanzar este nivel tan pronto como fuera posible. Por otra parte, como se quiere cubrir los casos $n = 2$ y $n = 3$, por lo menos, al usar n para denotar cualquiera de estos números se evitan repeticiones muy tediosas.

El primer capítulo persigue varios propósitos. Primero, y ante todo, establecer la conexión fundamental entre el álgebra lineal y la intuición geométrica. Ciertamente, hay cuando menos dos aspectos del álgebra lineal: la manipulación formal de los cálculos con matrices y la interpretación geométrica. No es mi deseo abogar por uno o por otro, pero sí considero que basar las manipulaciones formales en contextos geométricos proporciona un muy valioso antecedente a las personas que usan el álgebra lineal. Segundo, este primer capítulo brinda de inmediato ejemplos concretos, con coordenadas, de combinaciones lineales, perpendicularidad y otras nociones que se desarrollan más adelante. Además del contexto geométrico, el estudio de estas nociones suministra ejemplos de subespacios y también proporciona una interpretación fundamental de las ecuaciones lineales. Por tanto, el primer capítulo brinda una rápida visión global de muchos temas del libro. El contenido del primer capítulo también corresponde a los aspectos fundamentales de los cursos de cálculo en cuanto a funciones de

varias variables, lo cual nos permite desarrollar muchas cosas sin el empleo de las matrices más generales. Si en algún otro curso los estudiantes han cubierto el material correspondiente al Capítulo I, o si el instructor desea trabajar a fondo con las matrices recurriendo a otros medios, entonces puede omitirse el primer capítulo o usarse en forma selectiva para motivar a los alumnos o para aplicar ejemplos.

Después de este capítulo introductorio, comenzamos con ecuaciones lineales, matrices y la eliminación de Gauss. Este capítulo hace énfasis en los aspectos de cómputo del álgebra lineal. Luego trabajamos con espacios vectoriales, aplicaciones lineales y productos escalares, y sus relaciones con las matrices. Esto combina el aspecto computacional con el teórico.

Los determinantes se tratan en forma más breve que en la primera edición de este libro, y se suprimen varias pruebas. Los estudiantes interesados en la teoría pueden recurrir a un tratamiento más completo en libros teóricos de álgebra lineal.

Incluí un capítulo sobre valores propios y vectores propios, el cual permite practicar con nociones que se estudiaron previamente y sirve como introducción al material que se usa en forma constante en todas partes de las matemáticas y sus aplicaciones.

Agradezco efusivamente a Toby Orloff y Daniel Horn, quienes impartieron el curso con una versión preliminar de este libro, sus valiosos comentarios y correcciones.

Contenido

CAPÍTULO III

CAPÍTULO IV

CAPÍTULO V

CAPÍTULO VI

CAPÍTULO VII

CAPÍTULO I

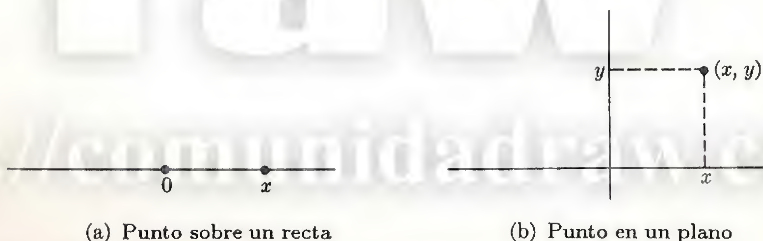
Vectores

El concepto de vector es básico para el estudio de las funciones de varias variables. Suministra una motivación geométrica para todo lo que sigue. Por tanto, las propiedades de los vectores, tanto algebraicas como geométricas, se estudiarán en detalle.

Una característica significativa de todos los enunciados y las demostraciones de esta parte es que no son ni más sencillos ni más difíciles de probar en el espacio de 3 dimensiones que lo que resultan ser en el espacio de 2 dimensiones.

I, §1. Definición de puntos en el espacio

Sabemos que se puede usar un número para representar un punto sobre una recta, una vez que se ha escogido una unidad de longitud.



(a) Punto sobre un recta (b) Punto en un plano

Figura 1

Se puede usar un par de números (esto es, una pareja de números) (x, y) para representar un punto en un plano.

Esto se puede representar como en la figura 1.

Observemos ahora que se puede usar una terna de números (x, y, z) para representar un punto en el espacio, esto es, en el espacio tridimensional o espacio de 3 dimensiones. Simplemente se introduce un eje más. La figura 2 ilustra este hecho.

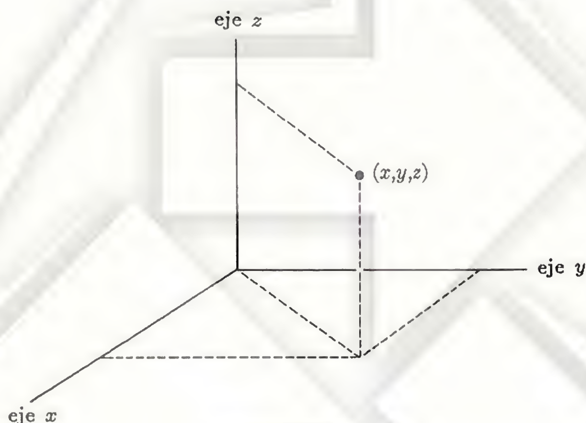


Figura 2

Podríamos usar también (x_1, x_2, x_3) en lugar de x, y, z. La recta se podría llamar espacio de dimensión 1 y el plano espacio de dimensión 2.

Así pues, podemos decir que un número representa un punto en el espacio de 1 dimensión. Una pareja representa un punto en el espacio de 2 dimensiones. Una terna representa un punto en el espacio de 3 dimensiones.

Aunque no podamos continuar dibujando diagramas, nada nos impide considerar una cuádrupla de números.

$$(x_1, x_2, x_3, x_4)$$

y afirmar que ésta es un punto en el espacio de 4 dimensiones. Una quíntupla sería un punto en el espacio de 5 dimensiones, luego vendría una séxtupla, una séptupla, una óctupla, ...

Podemos ir más allá y **definir un punto en el espacio de n dimensiones** como la n-tupla de números

$$(x_1, x_2, \ldots, x_n),$$

si n es un entero positivo. Denotaremos con una X mayúscula dicha n-tupla y procuraremos reservar las letras minúsculas para representar números y las letras mayúsculas para representar puntos. Llamaremos a los números $x_1, \ldots, x_n$ **coordenadas** del punto X. Por ejemplo, en el espacio de 3 dimensiones, 2

es la primera coordenada del punto $(2, 3, -4)$ y -4 es la tercera coordenada. Denotamos el espacio de n-dimensiones con $\mathbf{R}^n$.

En su mayoría, nuestros ejemplos ilustrarán el caso en que $n = 2$ o bien $n = 3$. Así, el lector podrá encontrar cualquiera de estos dos casos a lo largo del libro. Sin embargo, es necesario hacer tres comentarios.

Primero, tenemos que manejar $n = 2$ y $n = 3$, de manera que, con el objeto de evitar muchas repeticiones, resulta útil contar con una notación que cubra ambos casos simultáneamente, aun si con frecuencia repetimos la formulación de ciertos resultados por separado para ambos casos.

Segundo, ningún teorema o fórmula es más simple al suponer que $n = 2$ o $n = 3$.

Tercero, el caso $n = 4$ sí se presenta en física.

Ejemplo 1. Un ejemplo clásico del espacio de 3 dimensiones es, desde luego, el espacio en que vivimos. Después de haber escogido un origen y un sistema de coordenadas, podemos describir la posición de un punto (cuerpo, partícula, etc.) mediante 3 coordenadas. Además, como es bien sabido, resulta conveniente extender este espacio a otro de 4 dimensiones, en donde la cuarta coordenada es el tiempo, escogiendo el origen de éste en la fecha del nacimiento de Cristo, por ejemplo, aunque esto es puramente arbitrario (quizá sería más conveniente escoger como origen el nacimiento del sistema solar o el nacimiento de la Tierra, si pudiéramos determinarlos con precisión). Así, un punto con coordenada de tiempo negativa es un punto a.C. y un punto con coordenada de tiempo positiva es un punto d.C.

No obstante, no hay que pensar que "el tiempo es *la* cuarta dimensión". El espacio de 4 dimensiones que acabamos de mencionar es sólo un ejemplo posible. En economía, pongamos por caso, se usa un espacio muy diferente que toma como coordenadas, por ejemplo, la cantidad de dólares gastada en cada industria. De este modo, podríamos trabajar con un espacio de 7 dimensiones cuyas coordenadas correspondieran a las siguientes industrias:

1. Acero	2. Automóviles	3. Productos Agrícolas	4. Pesca
5. Productos químicos	6. Ropa	7. Transportes.	

Convengamos en que un megadólar por año es la unidad de medida. Luego, un punto

$$(1\,000, 800, 550, 300, 700, 200, 900)$$

en este espacio de 7 dimensiones significaría que la industria del acero gastó mil millones de dólares en el año indicado y que la industria química gastó 700 millones de dólares en ese año.

La idea de considerar el tiempo como una cuarta dimensión es muy antigua. Ya en la *Encyclopédie* de Diderot, que data del siglo XVIII, d'Alembert escribe lo siguiente en su artículo sobre "dimensión":

> Cette manière de considérer les quantités de plus de trois dimensions est aussi exacte que l'autre, car les lettres peuvent toujours être regardées comme représentant des nombres rationnels ou non. J'ai dit plus haut qu'il n'était pas possible de concevoir plus de trois dimensions. Un homme d'esprit de ma connaissance croit qu'on pourrait cependant regarder la durée comme une quatrième dimension, et que le produit temps par la solidité serait en quelque manière un produit de quatre dimensions; cette idée peut être contestée, mais elle a, ce me semble, quelque mérite, quand ce ne serait que celui de la nouveauté.
>
> *Encyclopédie,* Vol. 4 (1754), p. 1010

Lo que, traducido, dice:

> Esta manera de considerar cantidades que tienen más de tres dimensiones es tan correcta como la otra, debido a que siempre se puede considerar que las letras algebraicas representan números, ya sean racionales o no. Dije con anterioridad que no era posible concebir más de tres dimensiones. Sin embargo, un inteligente caballero que conozco piensa que se podría considerar a la duración como una cuarta dimensión y que el producto tiempo por volumen sería de algún modo un producto de cuatro dimensiones. Esta idea puede ponerse en tela de juicio pero, a mi entender, tiene algún mérito, aunque no sea sino el de ser novedosa.

Obsérvese que d'Alembert menciona a un "caballero inteligente" cuando se refiere, evidentemente, a sí mismo. Está proponiendo con cierta cautela lo que en esa época debe haber sido una idea incipiente; idea que en el siglo XX llega a ser bastante común.

D'Alembert también visualizó con claridad espacios de dimensión superior como "productos" de espacios de menor dimensión. Por ejemplo, podemos considerar un espacio de 3 dimensiones como si pusiéramos lado a lado las primeras dos coordenadas (x_1, x_2) y luego la tercera x_3. Por tanto, escribimos

$$\mathbf{R}^3 = \mathbf{R}^2 \times \mathbf{R}^1.$$

Usamos el signo del producto, que no deberá confundirse con otros "productos" como el producto de números. La palabra "producto" se emplea en dos contextos. En forma análoga, podemos escribir

$$\mathbf{R}^4 = \mathbf{R}^3 \times \mathbf{R}^1.$$

Hay otras maneras de expresar $\mathbf{R}^4$ como un producto, a saber,

$$\mathbf{R}^4 = \mathbf{R}^2 \times \mathbf{R}^2.$$

Esto significa que consideramos por separado las dos primeras coordenadas (x_1, x_2) y las dos últimas coordenadas (x_3, x_4). Más tarde retomaremos tales productos.

Ahora definiremos la adición de puntos. Si A y B son puntos, digamos en el espacio de 3 dimensiones,

$$A = (a_1, a_2, a_3) \quad \text{y} \quad B = (b_1, b_2, b_3)$$

entonces **definimos** $A + B$ como el punto cuyas coordenadas son

$$A + B = (a_1 + b_1, a_2 + b_2, a_3 + b_3).$$

Ejemplo 2. En el plano, si $A = (1,2)$ y $B = (-3,5)$, entonces

$$A + B = (-2,7).$$

En el espacio de 3 dimensiones, si $A = (-1, \pi, 3)$ y $B = (\sqrt{2}, 7, -2)$, entonces

$$A + B = (\sqrt{2} - 1, \pi + 7, 1).$$

Al usar una n neutral que cubriera ambos casos de los espacios de 2 y 3 dimensiones, los puntos se escribirían de la siguiente manera:

$$A = (a_1, \ldots, a_n), \qquad B = (b_1, \ldots, b_n),$$

y **definimos** $A + B$ como el punto cuyas coordenadas son

$$(a_1 + b_1, \ldots, a_n + b_n).$$

Observemos que se satisfacen las siguientes reglas:

1. $(A + B) + C = A + (B + C)$.
2. $A + B = B + A$.
3. Si suponemos que

$$O = (0, 0, \ldots, 0)$$

es el punto cuyas coordenadas son todas 0, entonces

$$O + A = A + O = A$$

para todo A.

4. Sea $A = (a_1, \ldots, a_n)$ y sea $-A = (-a_1, \ldots, -a_n)$. Entonces

$$A + (-A) = O.$$

Todas estas propiedades son muy sencillas, y son ciertas debido a que son ciertas para números y la adición de n-tuplas se define en términos de la adición de sus componentes, las cuales son números.

Nota. No se confunda el número 0 con la n-tupla $(0, \ldots, 0)$. Usualmente denotamos esta n-tupla con O, y también la llamamos cero, dado que en la práctica no puede ocurrir dificultad alguna.

Ahora interpretaremos geométricamente en el plano la adición y la multiplicación por números (en forma simultánea se puede visualizar lo que sucede en el espacio de 3 dimensiones).

Ejemplo 3. Sean $A = (2,3)$ y $B = (-1,1)$. Entonces

$$A + B = (1,4).$$

La figura se asemeja a un **paralelogramo** (Fig. 3).

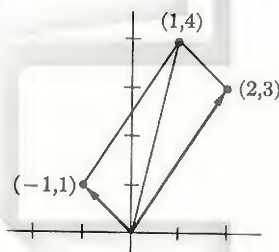


Figura 3

Ejemplo 4. Sean $A = (3,1)$ y $B = (1,2)$. Entonces

$$A + B = (4,3).$$

Vemos de nuevo que la representación geométrica de nuestra adición se asemeja a un **paralelogramo** (Fig. 4).

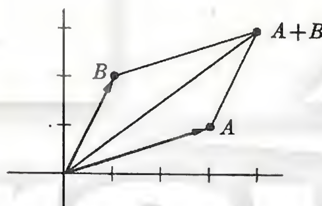


Figura 4

La razón por la cual la figura se asemeja a un **paralelogramo** se puede dar en términos de la geometría del plano de la manera siguiente. Obtenemos $B = (1,2)$ al comenzar en el origen $O = (0,0)$ y al movernos 1 unidad a la derecha y 2 hacia arriba. Para obtener $A+B$, comenzamos en A y nos movemos de nuevo 1 unidad a la derecha y 2 hacia arriba. Así, los segmentos de recta que se encuentran entre O y B y entre A y $A+B$ son las hipotenusas de los triángulos rectángulos cuyos catetos correspondientes son de la misma longitud y paralelos. Por consiguiente, los segmentos anteriores son paralelos y de la misma longitud, tal como se ilustra en la figura 5.

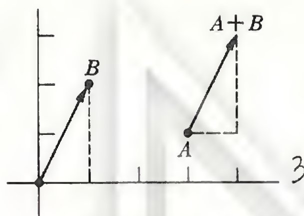


Figura 5

Ejemplo 5. Si $A = (3,1)$ de nuevo, entonces $-A = (-3,-1)$. Si marcamos este punto, vemos que $-A$ tiene sentido opuesto a A. Podemos considerar que $-A$ es la reflexión de A con respecto al origen.

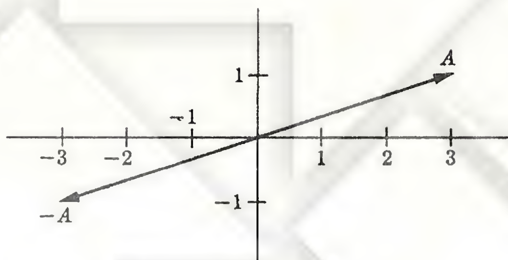


Figura 6

Ahora consideraremos la multiplicación de A por un número. Si c es cualquier número, definimos cA como el punto cuyas coordenadas son

$$(ca_1, \ldots, ca_n).$$

Ejemplo 6. Si $A = (2,-1,5)$ y $c = 7$, entonces $cA = (14,-7,35)$.

Es fácil verificar las siguientes reglas:

5. $c(A + B) = cA + cB$.
6. Si c_1 y c_2 son números, entonces

$$(c_1 + c_2)A = c_1A + c_2A \quad \text{y} \quad (c_1c_2)A = c_1(c_2A).$$

También nótese que

$$(-1)A = -A.$$

¿Cuál es la representación geométrica de la multiplicación por un número?

Ejemplo 7. Sean $A = (1,2)$ y $c = 3$. Entonces

$$cA = (3,6)$$

como se puede apreciar en la figura 7(a).

La multiplicación por 3 produce un alargamiento de A en 3 veces. En forma análoga, $\frac{1}{2}A$ da por resultado un acortamiento de A en $\frac{1}{2}$, es decir, A se reduce a la mitad de su tamaño. En general, si t es un número, $t > 0$, tA se interpreta como un punto en la misma dirección que A desde el origen, aunque a t veces la distancia. De hecho, definimos que A y B tienen la **misma dirección** si existe un número $c > 0$ tal que $A = cB$. Insistimos: esto significa que A y B tienen la misma dirección **con respecto al origen**. Para simplificar el lenguaje, omitimos las palabras "con respecto al origen".

La multiplicación por un número negativo invierte la dirección. Así, $-3A$ quedaría representado como en la figura 7(b).

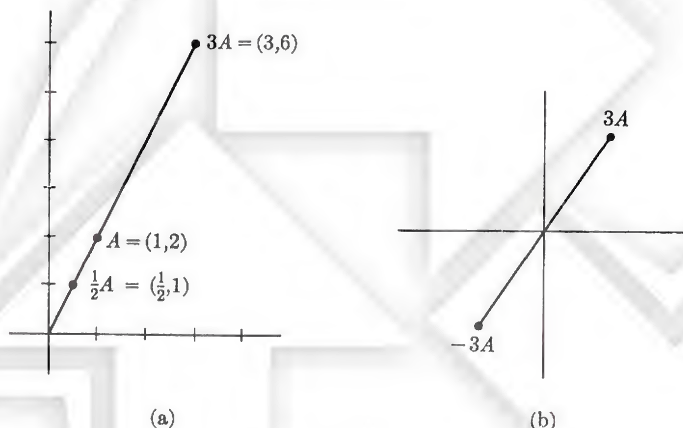


Figura 7

Decimos que dos vectores A y B (ninguno de los cuales es nulo) tienen **direcciones opuestas** si existe un número $c < 0$ tal que $cA = B$. Por lo cual, cuando $B = -A$, entonces A y B tienen direcciones opuestas.

Ejercicios I, §1

Encuentre $A + B$, $A - B$, $3A$, $-2B$ en cada uno de los siguientes casos. Marque los puntos que aparecen en los ejercicios 1 y 2 en una hoja de papel para graficar.

1. $A = (2, -1)$, $B = (-1, 1)$
2. $A = (-1, 3)$, $B = (0, 4)$
3. $A = (2, -1, 5)$, $B = (-1, 1, 1)$
4. $A = (-1, -2, 3)$, $B = (-1, 3, -4)$
5. $A = (\pi, 3, -1)$, $B = (2\pi, -3, 7)$
6. $A = (15, -2, 4)$, $B = (\pi, 3, -1)$
7. Sean $A = (1, 2)$ y $B = (3, 1)$. En una hoja de papel para graficar marque $A + B$, $A + 2B$, $A + 3B$, $A - B$, $A - 2B$, $A - 3B$.
8. Sean A y B como en el ejercicio 1. En una hoja de papel para graficar marque los puntos $A + 2B$, $A + 3B$, $A - 2B$, $A - 3B$ $A + \frac{1}{2}B$.

9. Considere A y B tal como se muestran en la figura 8. Marque el punto $A - B$, en cada caso.

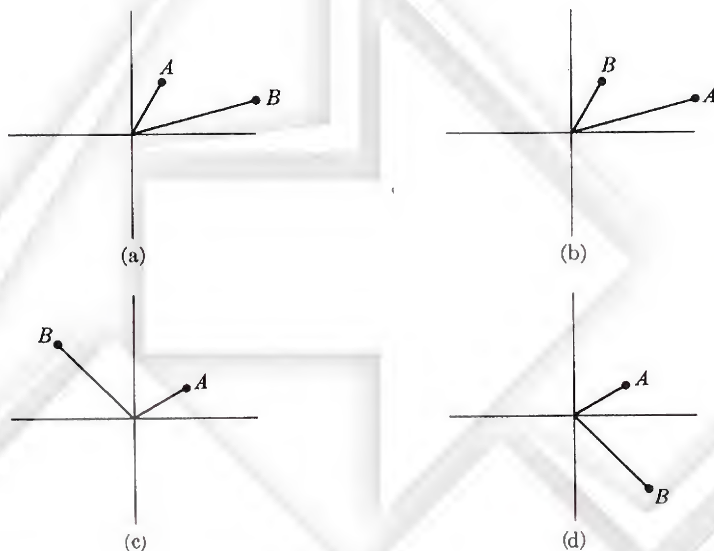


Figura 8

I, §2. Vectores anclados

Definimos un vector anclado como una pareja ordenada de puntos que representamos con $\overrightarrow{AB}$. (Éste *no* es un producto.) Podemos representarlo como una flecha que va de A a B. Designamos a A como **punto inicial** y a B como **punto final** del vector anclado (Fig. 9).

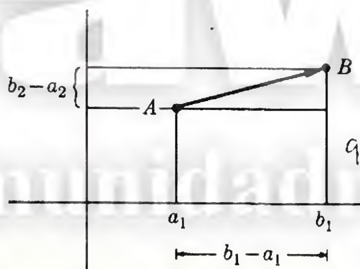


Figura 9

Observamos que, en el plano,

$$b_1 = a_1 + (b_1 - a_1).$$

De manera análoga,

$$b_2 = a_2 + (b_2 - a_2).$$

Esto significa que

$$B = A + (B - A).$$

Sean $\overrightarrow{AB}$ y $\overrightarrow{CD}$ dos vectores anclados. Decimos que son **equivalentes** si $B - A = D - C$. Todo vector anclado $\overrightarrow{AB}$ es equivalente a uno cuyo punto inicial es el origen, debido a que $\overrightarrow{AB}$ es equivalente a $\overrightarrow{O(B-A)}$. Es claro que éste es el único vector anclado cuyo punto inicial se encuentra en el origen y que es equivalente a $\overrightarrow{AB}$. Si representamos la ley del paralelogramo en el plano, entonces es evidente que la equivalencia de dos vectores anclados puede interpretarse geométricamente diciendo que las longitudes de los segmentos de recta determinados por la pareja de puntos son iguales y que las "direcciones" en que apuntan son las mismas.

En las siguientes figuras hemos dibujado los vectores anclados $\overrightarrow{O(B-A)}$, $\overrightarrow{AB}$ y $\overrightarrow{O(A-B)}$, $\overrightarrow{BA}$.

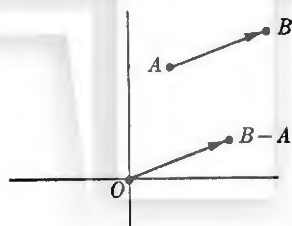


Figura 10

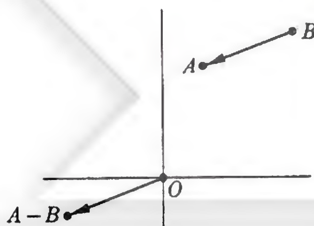


Figura 11

Ejemplo 1. Sean $P = (1, -1, 3)$ y $(Q) = (2, 4, 1)$. Entonces $\overrightarrow{PQ}$ es equivalente a $\overrightarrow{OC}$, donde $C = Q - P = (1, 5, -2)$. Si

$$A = (4, -2, 5) \quad \text{y} \quad B = (5, 3, 3),$$

entonces $\overrightarrow{PQ}$ es equivalente a $\overrightarrow{AB}$, ya que

$$Q - P = B - A = (1, 5, -2).$$

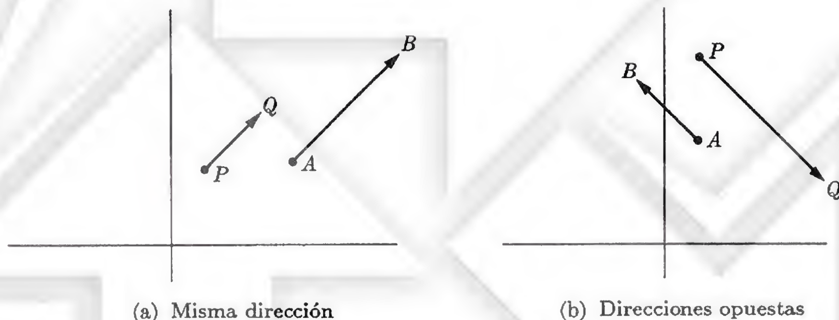
Dado un vector anclado $\overrightarrow{OC}$ cuyo punto inicial es el origen, decimos que se **encuentra anclado en el origen**. Dado cualqier vector anclado $\overrightarrow{AB}$, decimos que está **anclado en** A.

Un vector anclado en el origen se encuentra completamente determinado por su punto final. Con base en esto, designaremos como n-tupla a un punto o bien a un **vector**, dependiendo de la interpretación que tengamos en mente.

Se dice que dos vectores anclados $\overrightarrow{AB}$ y $\overrightarrow{PQ}$ son **paralelos** si existe un número $c \neq 0$ tal que $B - A = c(Q - P)$. Se dice que tienen la **misma dirección** si existe un número $c > 0$ tal que $B - A = c(Q - P)$, y se dice que tienen **direcciones opuestas** si existe un número $c < 0$ tal que

$$B - A = c(Q - P).$$

En las siguientes ilustraciones representamos dos casos de vectores anclados paralelos.



(a) Misma dirección (b) Direcciones opuestas

Figura 12

Ejemplo 2. Sean

$$P = (3,7) \quad \text{y} \quad Q = (-4,2).$$

Sean

$$A = (5,1) \quad \text{y} \quad B = (-16,-14).$$

Luego,

$$Q - P = (-7,-5) \quad \text{y} \quad B - A = (-21,-15).$$

En consecuencia $\overrightarrow{PQ}$ es paralelo a $\overrightarrow{AB}$, debido a que $B - A = 3(Q - P)$. Más aún, como $3 > 0$, vemos que $\overrightarrow{PQ}$ y $\overrightarrow{AB}$ tienen la misma dirección.

De manera semejante, cualquier definición enunciada para n-tuplas se puede aplicar a los vectores anclados. Por ejemplo, en la siguente sección definiremos lo que significa que las n-tuplas sean perpendiculares.

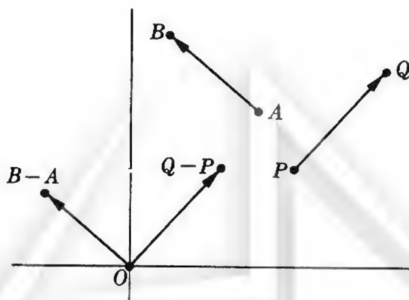


Figura 13

Entonces podemos decir que dos vectores anclados $\overrightarrow{AB}$ y $\overrightarrow{PQ}$ son **perpendiculares** entre sí, si $B - A$ es perpendicular a $Q - P$. En la figura 13 hemos hecho una representación de dichos vectores en el plano.

Ejercicios I, §2

En cada uno de los siguientes casos determine cuáles vectores anclados $\overrightarrow{PQ}$ y $\overrightarrow{AB}$ son equivalentes.

1. $P = (1, -1)$, $Q = (4, 3)$, $A = (-1, 5)$, $B = (5, 2)$.
2. $P = (1, 4)$, $Q = (-3, 5)$, $A = (5, 7)$, $B = (1, 8)$.
3. $P = (1, -1, 5)$, $Q = (-2, 3, -4)$, $A = (3, 1, 1)$, $B = (0, 5, 10)$.
4. $P = (2, 3, -4)$, $Q = (-1, 3, 5)$, $A = (-2, 3, -1)$, $B = (-5, 3, 8)$.

En cada uno de los siguientes casos determine cuáles vectores anclados $\overrightarrow{PQ}$ y $\overrightarrow{AB}$ son paralelos.

5. $P = (1, -1)$, $Q = (4, 3)$, $A = (-1, 5)$, $B = (7, 1)$.
6. $P = (1, 4)$, $Q = (-3, 5)$, $A = (5, 7)$, $B = (5, 6)$.
7. $P = (1, -1, 5)$, $Q = (-2, 3, -4)$, $A = (3, 1, 1)$, $B = (-3, 9, -17)$.
8. $P = (2, 3, -4)$, $Q = (-1, 3, 5)$, $A = (-2, 3, -1)$, $B = (-11, 3, -28)$.
9. Con el objeto de ilustrarlos, dibuje en una hoja de papel los vectores anclados de los ejercicios 1, 2, 5 y 6. Dibuje también los vectores anclados $\overrightarrow{QP}$ y $\overrightarrow{BA}$. Marque los puntos $Q - P$, $B - A$, $P - Q$ y $A - B$.

I, §3. Producto escalar

Convendremos en que, a lo largo de nuestro estudio, seleccionaremos vectores que siempre se encuentren en el mismo espacio n-dimensional. El lector puede restringirse sólo a los casos en que $n = 2$ y $n = 3$.

Sean $A = (a_1, a_2)$ y $B = (b_1, b_2)$ dos vectores en el espacio de dos dimensiones. Definimos su **producto escalar** como

$$A \cdot B = a_1 b_1 + a_2 b_2.$$

Sean $A = (a_1, a_2, a_3)$ y $B = (b_1, b_2, b_3)$ dos vectores en el espacio de tres dimensiones. Definimos su **producto escalar** de la siguiente manera:

$$A \cdot B = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

Sean $A = (a_1, \ldots, a_n)$ y $B = (b_1, \ldots, b_n)$ dos vectores en el espacio de n dimensiones, lo que cubre los dos casos anteriores con una sola notación. Definimos su **producto escalar** o **producto interior** $A \cdot B$ como sigue

$$a_1 b_1 + \cdots + a_n b_n.$$

Este producto es un **número**. Por ejemplo, si

$$A = (1, 3, -2) \qquad \text{y} \qquad B = (-1, 4, -3),$$

entonces

$$A \cdot B = -1 + 12 + 6 = 17.$$

Por el momento no daremos ninguna interpretación geométrica de este producto escalar; lo haremos posteriormente. Antes, vamos a deducir algunas propiedades importantes. Las básicas son:

PE 1. *Tenemos que* $A \cdot B = B \cdot A$.

PE 2. *Si* A, B *y* C *son tres vectores, entonces*

$$A \cdot (B + C) = A \cdot B + A \cdot C = (B + C) \cdot A.$$

PE 3. *Si* x *es un número, entonces*

$$(xA) \cdot B = x(A \cdot B) \qquad y \qquad A \cdot (xB) = x(A \cdot B).$$

PE 4. *Si* $A = O$ *es el vector nulo, entonces* $A \cdot A = 0$ *y, si no lo es, entonces*

$$A \cdot A > 0.$$

A continuación probamos estas propiedades.

Con respecto a la primera, tenemos que

$$a_1 b_1 + \cdots + a_n b_n = b_1 a_1 + \cdots + b_n a_n,$$

dado que, para dos números cualesquiera a y b, tenemos que $ab = ba$, lo que prueba la primera propiedad.

Para probar **PE 2**, sea $C = (c_1, \ldots, c_n)$. Entonces

$$B + C = (b_1 + c_1, \ldots, bn + c_n)$$

y

$$\begin{aligned} A \cdot (B + C) &= a_1(b_1 + c_1) + \cdots + a_n(b_n + c_n) \\ &= a_1 b_1 + a_1 c_1 + \cdots + a_n b_n + a_n c_n. \end{aligned}$$

Al reordenar los términos se obtiene

$$a_1 b_1 + \cdots + a_n b_n + a_1 c_1 + \cdots + a_n c_n,$$

que no es otra cosa que $A \cdot B + A \cdot C$. Esto prueba lo que queríamos. La propiedad **PE 3** se deja como ejercicio.

Por último, para probar **PE 4**, observamos que, si una coordenada a_i de A no es igual a 0, entonces existe un término $a_i^2 \neq 0$ y $a_i^2 > 0$ en el producto escalar

$$A \cdot A = a_1^2 + \cdots + a_n^2.$$

Puesto que todo término es ≥ 0, se infiere que la suma es > 0, tal como se quería demostrar.

En una gran parte del trabajo que realizaremos con los vectores, sólo usaremos las propiedades comunes de la adición, de la multiplicación por números y las cuatro propiedades del producto escalar. Más adelante los estudiaremos de manera formal. Por el momento, observe que existen otros objetos con los cuales el lector está familiarizado y que se pueden sumar, restar y multiplicar por números, por ejemplo, las funciones continuas definidas en un intervalo $[a, b]$.

Será conveniente escribir A^2 en vez de $A \cdot A$ para representar el producto escalar de un vector consigo mismo. (Éste es el único ejemplo en el que nos permitimos usar tal notación. Así pues, A^3 no tiene significado alguno.) Como ejercicio, verifique que se cumplen las siguientes identidades:

$$(A + B)^2 = A^2 + 2A \cdot B + B^2,$$

$$(A - B)^2 = A^2 - 2A \cdot B + B^2.$$

Un producto interior $A \cdot B$ puede muy bien ser igual a cero sin que ninguno de los vectores A o B sea el vector nulo. Por ejemplo, sean

$$A = (1, 2, 3) \quad \text{y} \quad B = (2, 1, -\tfrac{4}{3}).$$

Entonces,

$$A \cdot B = 0$$

Decimos que dos vectores A y B son **perpendiculares** entre sí (u **ortogonales**, como también les llamaremos), si $A \cdot B = 0$. Por ahora no es evidente que, en el plano, esta definición coincide con nuestra notación geométrica intuitiva de perpendicularidad. En la siguiente sección convenceremos al lector de tal coincidencia; en esta parte sólo veremos un ejemplo. Consideremos en $\mathbf{R}^3$ los tres vectores unitarios

$$E_1 = (1, 0, 0), \qquad E_2 = (0, 1, 0), \qquad E_3 = (0, 0, 1)$$

tal como se muestra en el diagrama (Fig. 14).

Veamos entonces que $E_1 \cdot E_2 = 0$ y que, en forma análoga, $E_i \cdot E_j = 0$ si $i \neq j$. Y así, estos vectores se consideran perpendiculares entre sí. Si $A = (a_1, a_2, a_3)$, entonces observamos que la i-ésima componente de A, a saber,

$$a_i = A \cdot E_i$$

es el producto interior de A con el i-ésimo vector unitario.. Advertimos que A es perpendicular a E_i (conforme a nuestra definición de perperdicularidad con el producto interior) si, y sólo si, su i-ésima componente es igual a 0.

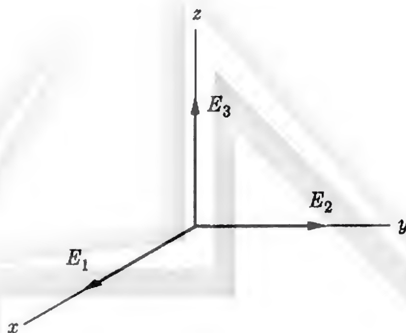


Figura 14

Ejercicios I, §3

1. Para cada una de las n-tuplas siguientes encuentre $A \cdot A$.
   (a) $A = (2, -1)$, $B = (-1, 1)$ (b) $A = (-1, 3)$, $B = (0, 4)$
   (c) $A = (2, -1, 5)$, $B = (-1, 1, 1)$ (d) $A = (-1, -2, 3)$, $B = (-1, 3, -4)$
   (e) $A = (\pi, 3, -1)$, $B = (2\pi, -3, 7)$ (f) $A = (15, -2, 4)$, $B = (\pi, 3, -1)$

2. Encuentre $A \cdot B$ para las n-tuplas del ejercicio anterior.

3. Usando sólo las cuatro propiedades del producto escalar, verifique con detalle las identidades dadas en el texto para $(A + B)^2$ y $(A - B)^2$.

4. ¿Cuáles de las siguientes parejas de vectores son perpendiculares entre sí?
   (a) $(1, -1, 1)$ y $(2, 1, 5)$ (b) $(1, -1, 1)$ y $(2, 3, 1)$
   (c) $(-5, 2, 7)$ y $(3, -1, 2)$ (d) $(\pi, 2, 1)$ y $(2, -\pi, 0)$

5. Sea A un vector perpendicular a todo vector X. Demuestre que $A = O$.

I, §4. La norma de un vector

Definimos como **norma** de un vector A, que denotamos con $\|A\|$, al número

$$\|A\| = \sqrt{A \cdot A}.$$

Como $A \cdot A \geq 0$, podemos extraer la raíz cuadrada. Algunas veces la norma también se denomina **magnitud** de A.

Cuando $n = 2$ y $A = (a, b)$, entonces

$$\|A\| = \sqrt{a^2 + b^2},$$

como se puede apreciar en la siguiente representación (Fig. 15).

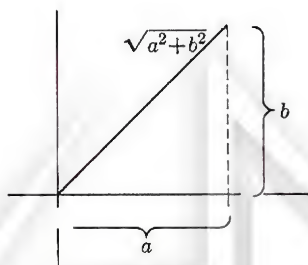


Figura 15

Ejemplo 1. Si $A = (1,2)$, entonces

$$\|A\| = \sqrt{1+4} = \sqrt{5}.$$

Cuando $n = 3$ y $A = (a_1, a_2, a_3)$, entonces

$$\|A\| = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

Ejemplo 2. Si $A = (-1,2,3)$, entonces

$$\|A\| = \sqrt{1+4+9} = \sqrt{14}.$$

Si $n = 3$, entonces la representación es como la de la figura 16, donde $A = (x,y,z)$.

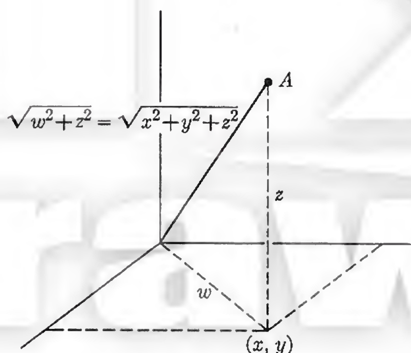


Figura 16

Si nos fijamos primero en las dos componentes (x,y), entonces, como se indicó, la longitud del segmento que une a $(0,0)$ con (x,y) es igual a $w = \sqrt{x^2+y^2}$.

Entonces, de nuevo, la norma de A sería, por el teorema de Pitágoras,

$$\sqrt{w^2+z^2} = \sqrt{x^2+y^2+z^2}.$$

Así, cuando $n = 3$, nuestra definición de norma es compatible con la geometría del teorema de Pitágoras.

En términos de coordenadas, $A = (a_1, \ldots, a_n)$ y vemos que

$$\|A\| = \sqrt{a_1^2 + \cdots + a_n^2}.$$

Si $A \neq O$, entonces $\|A\| \neq 0$, debido a que alguna coordenada $a_i \neq 0$, de manera que $a_i^2 > 0$ y, en consecuencia, $a_1^2 + \cdots + a_n^2 > 0$, por lo que $\|A\| \neq 0$.

Observe que, para cualquier vector A, tenemos que

$$\boxed{\|A| = \| - A\|.}$$

Esto se debe al hecho de que

$$(-a_1)^2 + \cdots + (-a_n)^2 = a_1^2 + \cdots + a_n^2,$$

ya que $(-1)^2 = 1$. Por supuesto, esto debe ser así, según la figura:

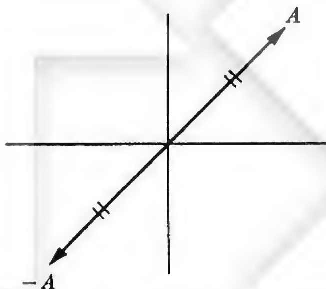


Figura 17

Recordemos que A y $-A$ tienen **direcciones opuestas**. Sin embargo, tienen la misma norma (magnitud, como algunas veces se dice cuando se habla de vectores).

Sean A y B dos puntos. Definimos la **distancia** entre A y B de la manera siguiente:

$$\|A - B\| = \sqrt{(A - B) \cdot (A - B)}.$$

Esta definición coincide con nuestra intuición geométrica cuando A y B son puntos en el plano (Fig. 18). Es lo mismo que la longitud del vector anclado $\overrightarrow{AB}$ o del vector anclado $\overrightarrow{BA}$.

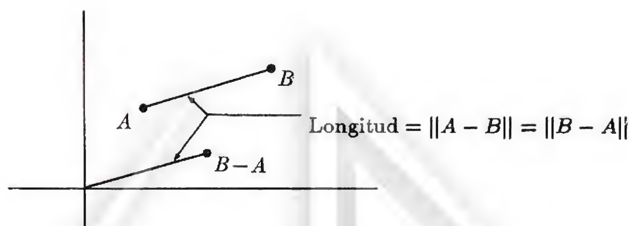


Figura 18

Ejemplo 3. Sean $A = (-1, 2)$ y $B = (3, 4)$. Entonces la longitud del vector anclado $\overrightarrow{AB}$ es igual a $\|B - A\|$. Pero $B - A = (4, 2)$. Por tanto,

$$\|B - A\| = \sqrt{16 + 4} = \sqrt{20}.$$

En la figura vemos que el lado horizontal tiene una longitud igual a 4 y el lado vertical tiene una longitud igual a 2, por lo que nuestras definiciones reflejan nuestra intuición geométrica derivada de Pitágoras.

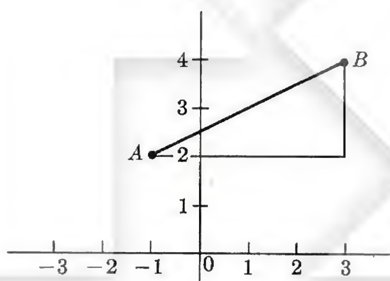


Figura 19

Sea P un punto en el plano y sea a un número > 0. El conjunto de los puntos X tales que

$$\|X - P\| < a$$

será denominado **disco abierto** de radio a con centro en P. El conjunto de los puntos X tales que

$$\|X - P\| \leq a$$

será denominado **disco cerrado** de radio a y centro P. El conjunto de los puntos X tales que

$$\|X - P\| = a$$

se conoce como círculo de radio a y centro P. En la figura 20 aparecen las ilustraciones de estos conceptos.

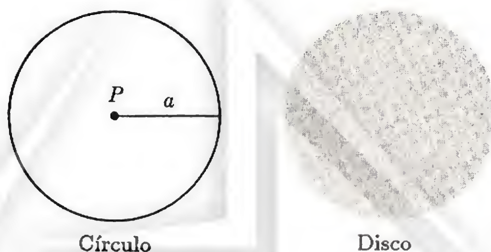


Figura 20

En el espacio de 3 dimensiones, el conjunto de los puntos X tales que

$$\|X - P\| < a$$

se denomina **bola abierta** de radio a y centro P. Al conjunto de los puntos X tales que

$$\|X - P\| \leq a$$

se le conoce como **bola cerrada** de radio a y centro P. El conjunto de los puntos X tales que

$$\|X - P\| = a$$

se llama **esfera** de radio a y centro P. En espacios de dimensión mayor se continúa usando la misma terminología de bola y esfera.

En la figura 21 se ilustran una esfera y una bola en el espacio de 3 dimensiones.

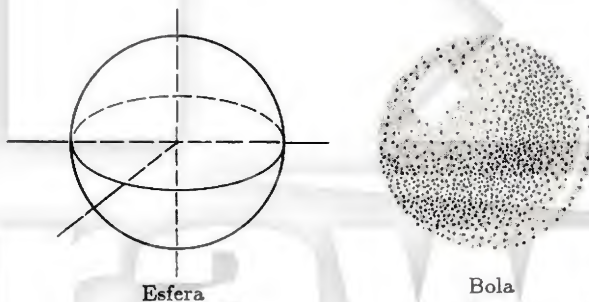


Figura 21

La esfera es la cáscara y la bola es la región que se encuentra dentro de la cáscara. La bola abierta consiste en la región situada dentro de la cáscara sin incluir la cáscara en sí. La bola cerrada está formada por la región contenida dentro de la cáscara y por la cáscara misma.

A partir de la geometría de la situación, también resulta razonable esperar que, si $c > 0$, entonces $\|A\| = c\|A\|$, es decir, si alargamos un vector A al multiplicarlo por un número positivo c, entonces la longitud también se incre-

menta en la misma cantidad. Vamos a verificar formalmente lo dicho, usando la definición de longitud.

Teorema 4.1. *Sea x un número. Entonces*

$$\|xA\| = |x|\,\|A\|$$

(el valor absoluto de x por la norma de A).

Demostración. Por definición, tenemos que

$$\|xA\|^2 = (xA) \cdot (xA),$$

que es igual a

$$x^2(A \cdot A)$$

debido a las propiedades del producto escalar. Al extraer la raíz cuadrada se obtiene lo que queríamos.

Sea S_1 la esfera de radio 1, con centro en el origen. Sea a un número > 0. Si x es un punto de la esfera S_1, entonces aX es un punto de la esfera de radio a, debido a que

$$\|aX\| = a\|X\| = a.$$

De esta manera, obtenemos todos los puntos de la esfera de radio a. (¿Demostración?) Por tanto, la esfera de radio a se obtiene al estirar la esfera de radio 1, mediante la multiplicación por a.

Una observación semejante se aplica a las bolas abierta y cerrada de radios a, las cuales se obtienen a partir de las bolas abierta y cerrada de radios 1, mediante la multiplicación por a.

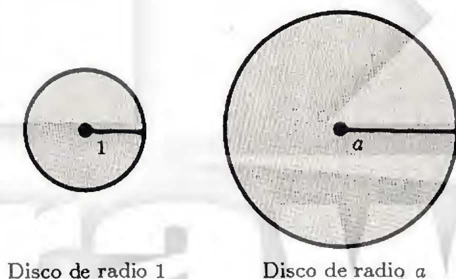


Figura 22

Decimos que un vector E es un vector **unitario** si $\|E\| = 1$. Dado cualquier vector A, sea $a = \|A\|$. Si $a \neq 0$, entonces

$$\frac{1}{a}A$$

es un vector unitario, dado que

$$\left\| \frac{1}{a}A \right\| = \frac{1}{a}a = 1.$$

Decimos que dos vectores A y B (ninguno de los cuales es el vector O) tienen la **misma dirección** si existe un número $c > 0$ tal que $cA = B$. En vista de esta definición, vemos que el vector

$$\frac{1}{\|A\|}A$$

es un vector unitario en la dirección de A (considerando que $A \neq O$).

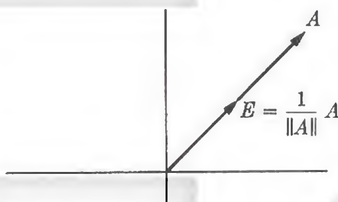


Figura 23

Si E es un vector unitario en la dirección de A y $\|A\| = a$, entonces

$$A = aE.$$

Ejemplo 4. Sea $A = (1, 2, -3)$. Entonces $\|A\| = \sqrt{14}$. En consecuencia, el vector unitario en la dirección de A es el vector

$$E = \left(\frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}}, \frac{-3}{\sqrt{14}}\right).$$

Advertencia. Hay tantos vectores unitarios como direcciones. Los tres **vectores unitarios canónicos** en el espacio de 3 dimensiones, a saber,

$$E_1 = (1, 0, 0), \qquad E_2 = (0, 1, 0), \qquad E_3 = (0, 0, 1)$$

son simplemente los tres vectores unitarios en las direcciones de los ejes de coordenadas.

También nos encontramos en posición de poder justificar nuestra definición de perpendicularidad. Dados A y B en el plano, la condición

$$\|A + B\| = \|A - B\|$$

(ilustrada en la figura 24(b)) coincide con la propiedad geométrica de que A debe ser perpendicular a B.

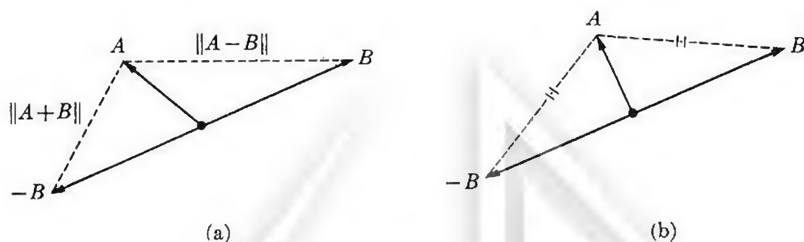


Figura 24

Probaremos que:

$$||A+B|| = ||A-B|| \text{ si, y sólo si, } A\cdot B = 0.$$

Denotemos con el símbolo $\Longleftrightarrow$ la expresión "si, y sólo si,". Entonces,

$$\begin{aligned} ||A+B|| = ||A-B|| &\Longleftrightarrow ||A+B||^2 = ||A-B||^2 \\ &\Longleftrightarrow A^2 + 2A\cdot B + B^2 = A^2 - 2A\cdot B + B^2 \\ &\Longleftrightarrow 4A\cdot B = 0 \\ &\Longleftrightarrow A\cdot B = 0. \end{aligned}$$

Esto prueba lo que queríamos.

Teorema de Pitágoras generalizado. *Si A y B son perpendiculares entre sí, entonces*

$$||A+B||^2 = ||A||^2 + ||B||^2.$$

En la figura 25 se ilustra el teorema.

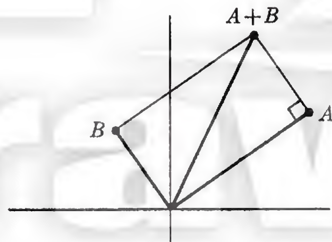


Figura 25

Para probarlo, usamos las definiciones, a saber:

$$\begin{aligned} ||A+B||^2 &= (A+B)\cdot(A+B) = A^2 + 2A\cdot B + B^2 \\ &= ||A||^2 + ||B||^2, \end{aligned}$$

ya que, por definición, $A\cdot B = 0$ y $A\cdot A = ||A||^2$, $B\cdot B = ||B||^2$.

Observación. Si A es perpendicular a B y x es cualquier número, entonces A también es perpendicular a xB, debido a que

$$A \cdot xB = xA \cdot B = 0.$$

Ahora emplearemos la noción de perpendicularidad para deducir la noción de **proyección.** Sean A y B dos vectores y $B \neq 0$. Sea P el punto sobre la recta que contiene a $\overrightarrow{OB}$, tal que $\overrightarrow{PA}$ es perpendicular a $\overrightarrow{OB}$, como se muestra en la figura 26(a).

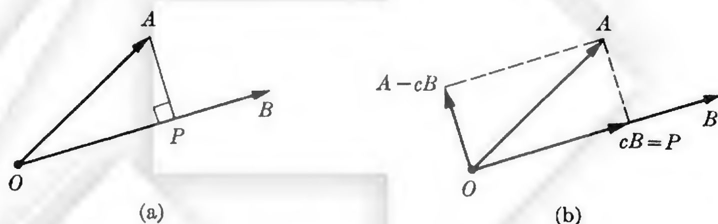


Figura 26

Podemos escribir

$$P = cB$$

para algún número c. Queremos encontrar este número c de forma explícita en términos de A y B. La condición $\overrightarrow{PA} \perp \overrightarrow{OB}$ significa que

$$A - P \text{ es perpendicular a } B,$$

y como $P = cB$, esto significa que

$$(A - cB) \cdot B = 0,$$

en otras palabras,

$$A \cdot B - cB \cdot B = 0.$$

Podemos despejar c y encontramos que $A \cdot B = cB \cdot B$, de manera que

$$\boxed{c = \frac{A \cdot B}{B \cdot B}.}$$

Recíprocamente, si tomamos este valor de c y luego usamos la distributividad, al multiplicar escalarmente a $A - cB$ por B se obtiene 0, de manera que $A - cB$ es perpendicular a B. En consecuencia, hemos visto que existe un único número c tal que $A - cB$ es perpendicular a B, y c está dado por la fórmula anterior.

Definición. La **componente** de A a lo largo de B es el número $c = \dfrac{A \cdot B}{B \cdot B}$.

La **proyección** de A a lo largo de B es el vector $cB = \dfrac{A \cdot B}{B \cdot B} B$.

Ejemplo 5. Supongamos que

$$B = E_i = (0, \ldots, 0, 1, 0, \ldots, 0)$$

es el i-ésimo vector unitario, donde 1 está en la i-ésima componente y 0 en todas las otras componentes.

$$\textit{Si } A = (a_1, \ldots, a_n), \ \textit{entonces } A \cdot E_i = a_i.$$

Por lo que $A \cdot E_i$ es la i-ésima componente ordinaria de A.

En forma más general, si B es un vector unitario, no necesariamente uno de los E_i, entonces simplemente tenemos que

$$c = A \cdot B$$

ya que $B \cdot B = 1$, por definición de vector unitario.

Ejemplo 6. Sean $A = (1, 2, -3)$ y $B = (1, 1, 2)$. Entonces la componente de A a lo largo de B es el **número**

$$c = \frac{A \cdot B}{B \cdot B} = \frac{-3}{6} = -\frac{1}{2}.$$

En consecuencia, la proyección de A a lo largo de B es el **vector**

$$cB = (-\tfrac{1}{2}, -\tfrac{1}{2}, -1).$$

Nuestra construcción brinda una interpretación geométrica inmediata del producto escalar. A saber, supongamos que $A \neq O$ y observemos el ángulo θ formado entre A y B (Fig. 27). Entonces, según la geometría del plano, vemos que

$$\cos\theta = \frac{c\|B\|}{\|A\|},$$

o bien, al sustituir el valor de c obtenido con anterioridad,

$$\boxed{A \cdot B = \|A\|\,\|B\|\cos\theta} \quad \text{y} \quad \boxed{\cos\theta = \frac{A \cdot B}{\|A\|\,\|B\|}.}$$

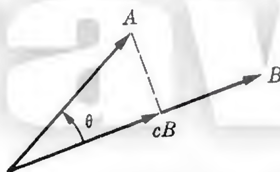


Figura 27

En algunos tratados sobre vectores, se considera la relación

$$A \cdot B = \|A\|\,\|B\|\cos\theta$$

como definición del producto escalar. Ésta se encuentra sujeta a las siguientes desventajas, por no decir objeciones:

(a) Las cuatro propiedades del producto escalar **PE 1** hasta **PE 4** de ninguna manera son obvias.
(b) Aun en el espacio de 3 dimensiones, se tiene que confiar en la intuición geométrica para obtener el coseno del ángulo formado entre A y B, y esta intuición resulta menos clara que en el plano. En espacios de dimensión mayor falla aún más.
(c) Resulta extremadamente difícil trabajar con tal definición para lograr propiedades adicionales del producto escalar.

Por tal motivo preferimos establecer fundamentos algebraicos obvios y después recuperar con gran sencillez todas las propiedades. Empleamos geometría del plano para ver la expresión

$$A \cdot B = \|A\| \, \|B\| \cos \theta.$$

Después de estudiar algunos ejemplos, probaremos la desigualdad que nos permite justificar esta fórmula en el espacio de n dimensiones.

Ejemplo 7. Sean $A = (1, 2, -3)$ y $B = (2, 1, 5)$. Encuentre el coseno del ángulo θ formado entre A y B.

Por definición,

$$\cos \theta = \frac{A \cdot B}{\|A\| \, \|B\|} = \frac{2 + 2 - 15}{\sqrt{14}\sqrt{30}} = \frac{-11}{\sqrt{420}}.$$

Ejemplo 8. Encuentre el coseno del ángulo formado entre los dos vectores anclados $\overrightarrow{PQ}$ y $\overrightarrow{PR}$, donde

$$P = (1, 2, -3), \qquad Q = (-2, 1, 5), \qquad R = (1, 1, -4).$$

La representación es como la siguiente:

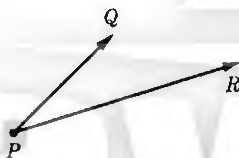


Figura 28

Hagamos

$$A = Q - P = (-3, -1, 8) \qquad \text{y} \qquad B = R - P = (0, -1, -1).$$

Entonces el ángulo formado entre $\overrightarrow{PQ}$ y $\overrightarrow{PR}$ es el mismo que el formado entre A y B. Por tanto, su coseno es igual a

$$\cos \theta = \frac{A \cdot B}{\|A\| \, \|B\|} = \frac{0 + 1 - 8}{\sqrt{74}\sqrt{2}} = \frac{-7}{\sqrt{74}\sqrt{2}}.$$

Probaremos propiedades adicionales de la norma y del producto escalar usando nuestros resultados sobre perpendicularidad. Obsérvese primero un caso especial. Si

$$E_i = (0, \dots, 0, 1, 0, \dots, 0)$$

es el i-ésimo vector unitario de $\mathbf{R}^n$, y

$$A = (a_1, \dots, a_n),$$

entonces

$$A \cdot E_i = a_i$$

es la i-ésima componente de A, esto es, la componente de A a lo largo de E_i. Tenemos que

$$|a_i| = \sqrt{a_i^2} \leq \sqrt{a_1^2 + \cdots + a_n^2} = ||A||,$$

de manera que el valor absoluto de cada componente de A es igual, a lo más, a la longitud de A.

No tenemos que trabajar, como lo acabamos de hacer, solamente con el vector unitario especial. Sea E cualquier vector unitario, es decir, un vector cuya norma sea igual a 1. Sea c la componente de A a lo largo de E. Vimos que

$$c = A \cdot E.$$

Luego, $A - cE$ es perpendicular a E, y

$$A = A - cE + cE.$$

Entonces $A - cE$ también es perpendicular a cE y, por el teorema de Pitágoras, encontramos que

$$||A||^2 = ||A - cE||^2 + ||cE||^2 = ||A - cE||^2 + c^2.$$

Así tenemos las desigualdades $c^2 \leq ||A||^2$ y $|c| \leq ||A||$.

En el siguiente teorema generalizamos esta desigualdad para un producto escalar $A \cdot B$ para el caso en que B no es necesariamente un vector unitario.

Teorema 4.2. *Sean A y B dos vectores en $\mathbf{R}^n$. Entonces*

$$|A \cdot B| \leq ||A|| \, ||B||.$$

Demostración. Si $B = O$, entonces ambos miembros de la desigualdad son iguales a 0, de modo que nuestra afirmación es obvia. Supongamos que $B \neq 0$. Sea c la componente de A a lo largo de B, de manera que $c = (A \cdot B)/(B \cdot B)$. Escribamos

$$A = A - cB + cB.$$

Por Pitágoras,

$$||A||^2 = ||A - cB||^2 + ||cB||^2 = ||A - cB||^2 + c^2||B||^2.$$

Por tanto, $c^2||B||^2 \leq ||A||^2$. Pero

$$c^2||B||^2 = \frac{(A \cdot B)^2}{(B \cdot B)^2}||B||^2 = \frac{|A \cdot B|^2}{||B||^4}||B||^2 = \frac{|A \cdot B|^2}{||B||^2}.$$

Por consiguiente,

$$\frac{|A\cdot B|^2}{\|B\|^2} \le \|A\|^2.$$

Para concluir la demostración, multiplicamos por $\|B\|^2$ y extraemos la raíz cuadrada.

Debido al teorema 4.2, vemos que, para los vectores A y B en el espacio de n dimensiones, el número

$$\frac{A\cdot B}{\|A\|\,\|B\|}$$

tiene valor absoluto ≤ 1. En consecuencia

$$-1 \le \frac{A\cdot B}{\|A\|\,\|B\|} \le 1,$$

y existe un ángulo θ único tal que $0 \le \theta \le \pi$, y tal que

$$\cos\theta = \frac{A\cdot B}{\|A\|\,\|B\|}.$$

Definimos este ángulo como el **ángulo formado por A y B**.

A la desigualdad del teorema 4.2 se le conoce como **desigualdad de Schwarz**.

Teorema 4.3. *Sean A y B vectores. Entonces*

$$\|A+B\| \le \|A\| + \|B\|.$$

Demostración. Ambos lados de esta desigualdad son positivos o iguales a 0. Por tanto, será suficiente probar que sus cuadrados satisfacen la desigualdad deseada; en otras palabras,

$$(A+B)\cdot(A+B) \le (\|A\| + B\|)^2.$$

Para lograrlo, consideremos

$$(A+B)\cdot(A+B) = A\cdot A + 2A\cdot B + B\cdot B.$$

En vista de nuestro resultado previo, esta relación satisface la desigualdad

$$\le \|A\|^2 + 2\|A\|\,\|B\| + \|B\|^2,$$

donde esta última expresión no es otra cosa que

$$(\|A\| + \|B\|)^2.$$

Nuestro teorema está probado.

El teorema 4.3 se conoce como **desigualdad del triángulo**. La razón para esto es que, si dibujamos un triángulo como el de la figura 29, entonces el teorema 4.3 expresa el hecho de que la longitud de un lado es $\le$ que la suma de las longitudes de los otros dos lados.

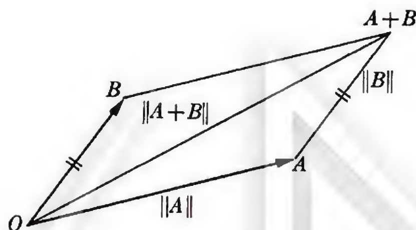


Figura 29

Observación. En ninguna de las demostraciones se hace uso de coordenadas; sólo se emplean las propiedades **PE 1** hasta **PE 4** del producto escalar. En consecuencia, siguen siendo válidas en situaciones más generales. Vea el Capítulo VI. En el espacio de dimensión n, nos dan desigualdades que de ninguna manera son obvias cuando se expresan en términos de coordenadas. Por ejemplo, la desigualdad de Schwarz, en términos de coordenadas, tiene el siguiente aspecto:

$$|a_1b_1 + \cdots + a_nb_n| \leq (a_1^2 + \cdots + a_n^2)^{1/2}(b_1^2 + \cdots + b_n^2)^{1/2}.$$

Sólo trate el lector de probar esta desigualdad directamente, sin usar la intuición "geométrica" de Pitágoras, y vea hasta dónde puede llegar.

Ejercicios I, §4

1. Encuentre la norma del vector A en los siguientes casos.
   (a) $A = (2, -1)$, $B = (-1, 1)$
   (b) $A = (-1, 3)$, $B = (0, 4)$
   (c) $A = (2, -1, 5)$, $B = (-1, 1, 1)$
   (d) $A = (-1, -2, 3)$, $B = (-1, 3, -4)$
   (e) $A = (\pi, 3, -1)$, $B = (2\pi, -3, 7)$
   (f) $A = (15, -2, 4)$, $B = (\pi, 3, -1)$

2. En los casos que abarca el ejercicio 1, encuentre la norma del vector B.

3. En los casos que abarca el ejercicio 1, encuentre la proyección de A a lo largo de B.

4. En los casos que abarca el ejercicio 1, encuentre la proyección de B a lo largo de A.

5. Encuentre el coseno del ángulo formado entre los vectores A y B.
   (a) $A = (1, -2)$ y $B = (5, 3)$
   (b) $A = (-3, 4)$ y $B = (2, -1)$
   (c) $A = (1, -2, 3)$ y $B = (-3, 1, 5)$
   (d) $A = (-2, 1, 4)$ y $B = (-1, -1, 3)$
   (e) $A = (-1, 1, 0)$ y $B = (2, 1, -1)$

6. Determine el coseno de los ángulos del triángulo cuyos vértices son
   (a) $(2,-1,1)$, $(1,-3,-5)$, $(3,-4,-4)$
   (b) $(3,1,1)$, $(-1,2,1)$, $(2,-2,5)$

7. Sean $A_1, \ldots, A_r$ vectores no nulos que son mutuamente perpendiculares, en otras palabras, $A_i \cdot A_j = 0$ si $i \neq j$. Sean $c_1, \cdots, c_r$ números tales que

$$c_1 A_1 + \cdots + c_r A_r = O.$$

   Demuestre que todos los $c_i = 0$.

8. Para cualesquiera vectores A y B, pruebe las siguientes relaciones:
   (a) $\|A+B\|^2 + \|A-B\|^2 = 2\|A\|^2 + 2\|B\|^2$
   (b) $\|A+B\|^2 = \|A\|^2 + \|B\|^2 + 2A \cdot B$
   (c) $\|A+B\|^2 - \|A-B\|^2 = 4A \cdot B$.
   Interprete (a) como una "ley del paralelogramo".

9. Demuestre que, si θ es el ángulo comprendido entre A y B, entonces

$$\|A-B\|^2 = \|A\|^2 + \|B\|^2 - 2\|A\|\,\|B\| \cos\theta.$$

10. Sean A, B y C tres vectores no nulos. Si $A \cdot B = A \cdot C$, muestre mediante un ejemplo que no necesariamente resulta que $B = C$.

I, §5. Rectas paramétricas

Definimos la ecuación **paramétrica** o **representación paramétrica** de una recta que pasa por un punto P en la dirección de un vector $A \neq 0$ de la siguiente manera:

$$X = P + tA,$$

donde t recorre todos los números (Fig. 30).

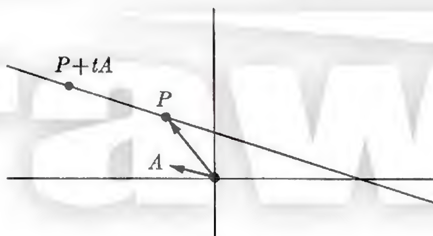


Figura 30

Al dar tal representación paramétrica, podemos imaginarnos un insecto que parte de un punto P en el tiempo $t = 0$ y se mueve en la dirección de A. En el tiempo t, el insecto se encuentra en la posición $P + tA$. Así podemos interpretar físicamente la representación paramétrica como una descripción del

movimiento, en la cual A se puede considerar como la velocidad del insecto. En un determinado tiempo t, el insecto se encuentra en el punto

$$X(t) = P + tA,$$

que es la **posición** del insecto en el tiempo t.

Esta representación paramétrica también es útil para describir el conjunto de puntos que están en el segmento de recta determinado por dos puntos dados. Sean P y Q dos puntos. Entonces el **segmento** determinado por P y Q consiste en todos los puntos

$$S(t) = P + t(Q - P) \qquad \text{donde} \qquad 0 \leq t \leq 1.$$

En realidad, $\overrightarrow{O(Q-P)}$ es un vector que tiene la misma dirección que $\overrightarrow{PQ}$, tal como se muestra en la figura 31.

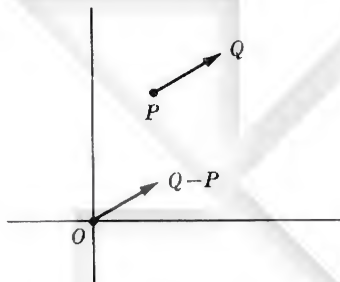


Figura 31

Cuando $t = 0$, tenemos que $S(0) = P$, de manera que, en el tiempo $t = 0$, el insecto se encuentra en P. Cuando $t = 1$ tenemos

$$S(1) = P + (Q - P) = Q,$$

de modo que, cuando $t = 1$, el insecto se encuentra en Q. Conforme t va de 0 a 1, el insecto avanza de P a Q.

Ejemplo 1. Sean $P = (1, -3, 4)$ y $Q = (5, 1, -2)$. Halle las coordenadas del punto que se encuentra a un tercio del camino de P a Q.

Sea $S(t)$, como se indicó con anterioridad, la representación paramétrica del segmento que une a P con Q. El punto deseado es $S(1/3)$, esto es:

$$S\left(\frac{1}{3}\right) = P + \frac{1}{3}(Q - P) = (1, -3, 4) + \frac{1}{3}(4, 4, -6)$$

$$= \left(\frac{7}{3}, \frac{-5}{3}, 2\right).$$

Advertencia. El punto deseado en el ejemplo que se acaba de ver *no* está dado por

$$\frac{P+Q}{3}.$$

Ejemplo 2. Encuentre una representación paramétrica para la recta que pasa por los dos puntos $P = (1, -3, 1)$ y $Q = (-2, 4, 5)$.

Primero tenemos que encontrar un vector en la dirección de la recta. Hagamos

$$A = P - Q,$$

por lo que

$$A = (3, -7, -4).$$

Por consiguiente, la representación paramétrica de la recta es

$$X(t) = P + tA = (1, -3, 1) + t(3, -7, -4).$$

Observación. También sería correcto dar una representación paramétrica de la recta de la siguiente manera:

$$Y(t) = P + tB \qquad \text{donde} \qquad B = Q - P.$$

Sin embargo, interpretadas en términos del movimiento del insecto, una parametrización da la posición de un insecto que se mueve en una dirección a lo largo de la recta, a partir de P en el tiempo $t = 0$, mientras que la otra parametrización da la posición de otro insecto que se mueve en la dirección opuesta a lo largo de la recta, partiendo también de P en el tiempo $t = 0$.

Ahora estudiaremos la relación que hay entre una representación paramétrica y la ecuación ordinaria de una recta en el plano.

Suponga que trabajamos en el plano y escriba las coordenadas de un punto X como (x, y). Sean $P = (p, q)$ y $A = (a, b)$. Entonces, en términos de las coordenadas, podemos escribir

$$x = p + ta, \qquad y = q + tb.$$

De modo que podemos eliminar t y obtener la ecuación usual que relaciona a x y y.

Ejemplo 3. Sean $P = (2, 1)$ y $A = (-1, 5)$. Entonces la representación paramétrica de la recta que pasa por P en la dirección de A nos da

(*) $$x = 2 - t, \qquad y = 1 + 5t.$$

Al multiplicar por 5 la primera ecuación y al sumar se obtiene

(**) $$5x + y = 11,$$

que es la conocida ecuación de una recta.

Esta eliminación de t muestra que toda pareja (x, y) que satisface la representación paramétrica $(*)$ para algún valor de t también satisface la ecuación $(**)$. Recíprocamente, suponga que tenemos una pareja de números (x, y) que satisface $(**)$. Sea $t = 2 - x$. Entonces

$$y = 11 - 5x = 11 - 5(2 - t) = 1 + 5t.$$

En consecuencia, existe algún valor de t que satisface la ecuación $(*)$. Por tanto, hemos probado que las parejas (x, y) que son soluciones de $(**)$ son exactamente las mismas parejas de números que aquellas que se obtienen al dar valores arbitrarios a t en $(*)$. Así, la recta se puede describir en forma paramétrica como en $(*)$ o en términos de su ecuación usual $(**)$. Al comenzar con la ecuación ordinaria

$$5x + y = 11,$$

hacemos $t = 2 - x$ con el objeto de recuperar la parametrización específica de $(*)$.

Cuando parametrizamos una recta en la forma

$$X = P + tA,$$

tenemos, por supuesto, una infinidad de posibilidades para escoger P en la recta y también una infinidad de posibilidades para escoger A, difiriendo en un múltiplo escalar. Siempre podemos seleccionar al menos uno. A saber, dada una ecuación

$$ax + by = c,$$

donde a, b y c son números, supongamos que $a \neq 0$. Usemos y como parámetro y hagamos

$$y = t.$$

Entonces podemos despejar x, a saber,

$$x = \frac{c}{a} - \frac{b}{a}t.$$

Sean $P = (c/a, 0)$ y $A = (-b/a, 1)$. Vemos que un punto arbitrario (x, y) que satisface la ecuación

$$ax + by = c$$

se puede expresar en forma paramétrica, a saber,

$$(x, y) = P + tA.$$

En dimensiones más grandes, al comenzar con una representación paramétrica

$$X = P + tA,$$

no podemos eliminar t *y, por tanto, la representación paramétrica es la única disponible para describir una recta.*

Ejercicios I, §5

1. Encuentre una representación paramétrica para la recta que pasa por las siguientes parejas de puntos.
   (a) $P_1 = (1, 3, -1)$ y $P_2 = (-4, 1, 2)$
   (b) $P_1 = (-1, 5, 3)$ y $P_2 = (-2, 4, 7)$

Encuentre una representación paramétrica para la recta que pasa por los siguientes puntos.

2. $(1, 1, -1)$ y $(-2, 1, 3)$
3. $(-1, 5, 2)$ y $(3, -4, 1)$
4. Sean $P = (1, 3, -1)$ y $Q = (-4, 5, 2)$. Determine las coordenadas de los siguientes puntos:
   (a) El punto medio del segmento de recta que une a P con Q.
   (b) Los dos puntos de este segmento de recta que se encuentran a un tercio y a dos tercios del camino de P a Q.
   (c) El punto que está a una quinta parte del camino de P a Q.
   (d) El punto que se encuentra a las dos quintas partes del camino de P a Q.
5. Si P y Q son dos puntos arbitrarios del espacio de n dimensiones, dé la fórmula general para el punto medio del segmento de recta determinado por P y Q.

I, §6. Planos

En el espacio de 3 dimensiones podemos describir planos mediante una ecuación análoga a la ecuación sencilla de la recta. Procedemos de la siguiente manera:

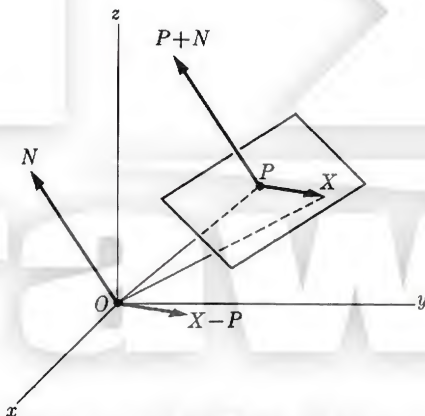


Figura 32

Sea P un punto en el espacio de 3 dimensiones y consideremos un vector anclado $\overrightarrow{ON}$. Definimos el **plano que pasa por P que es perpendicular a $\overrightarrow{ON}$** como la colección de todos los puntos X tales que el vector anclado $\overrightarrow{PX}$

es perpendicular a $\overrightarrow{ON}$. Conforme a nuestras definiciones, esto equivale a la condición

$$(X - P) \cdot N = 0,$$

que también se puede escribir como

$$\boxed{X \cdot N = P \cdot N.}$$

También diremos que éste es el plano perpendicular a N y que consta de todos los vectores X tales que $X - P$ es perpendicular a N. En la figura 32 hemos dibujado una situación típica en el espacio de 3 dimensiones.

Además de decir que N es **perpendicular** al plano, también se dice que N es **normal** al plano.

Sea t un número $\neq 0$. Entonces el conjunto de los puntos X tales que

$$(X - P) \cdot N = 0$$

coincide con el conjunto de los puntos X tales que

$$(X - P) \cdot tN = 0.$$

Por tanto, podemos decir que nuestro plano es aquel que pasa por P y es perpendicular a la **recta** en la dirección de N. Para hallar la ecuación del plano, podríamos usar cualquier vector tN (con $t \neq 0$) en lugar de N.

Ejemplo 1. Sean

$$P = (2, 1, -1) \qquad \text{y} \qquad N = (-1, 1, 3).$$

Sea $X = (x, y, z)$. Entonces

$$X \cdot N = (-1)x + y + 3z.$$

Por consiguiente, la ecuación del plano que pasa por P y es perpendicular a N es

$$-x + y + 3z = -2 + 1 - 3,$$

o bien,

$$-x + y + 3z = -4.$$

Observe que en el espacio de 2 dimensiones, donde $X = (x, y)$, las fórmulas conducen a la ecuación de la recta en el caso ordinario.

Ejemplo 2. La ecuación de la recta en el plano (x, y), que pasa por $(4, -3)$ y que es perpendicular a $(-5, 2)$, es

$$-5x + 2y = -20 - 6 = -26.$$

Nos encontramos ahora en posición de interpretar los coeficientes $(-5, 2)$ de x y y que aparecen en esta ecuación. Estos coeficientes dan origen a un vector perpendicular a la recta. **En cualquier ecuación**

$$ax + by = c$$

el vector (a, b) es perpendicular a la recta determinada por la ecuación. En forma análoga, en el espacio de 3 dimensiones, el vector (a, b, c) es perpendicular al plano determinado por la ecuación

$$ax + by + cz = d.$$

Ejemplo 3. El plano determinado por la ecuación

$$2x - y + 3z = 5$$

es perpendicular al vector $(2, -1, 3)$. Por supuesto, si queremos encontrar un punto en ese plano, tenemos muchas opciones. Podemos dar valores arbitrarios a x y a y y luego despejar z. Para obtener un punto en concreto, hagamos $x = 1$, y $y = 1$. Luego despejemos z, a saber,

$$3z = 5 - 2 + 1 = 4,$$

de manera que $z = \frac{4}{3}$. Así,

$$(1, 1, \tfrac{4}{3})$$

es un punto en el plano.

En el espacio de n dimensiones se dice que la ecuación $X \cdot N = P \cdot N$ es la que corresponde a un **hiperplano.** Por ejemplo,

$$3x - y + z + 2w = 5$$

es la ecuación de un hiperplano en el espacio de 4 dimensiones que es perpendicular a $(3, -1, 1, 2)$.

Se dice que dos vectores, A y B, son paralelos si existe un número $c \neq 0$ tal que $cA = B$. Se dice que dos rectas son **paralelas** si, dados dos puntos distintos entre sí P_1, Q_1 de la primera recta y P_2, Q_2 de la segunda, los vectores

$$P_1 - Q_1$$

y

$$P_2 - Q_2$$

son paralelos.

Se dice que dos planos son **paralelos** (en el espacio de 3 dimensiones) si sus vectores normales son paralelos. Se dice que son **perpendiculares** si sus vectores normales son perpendiculares. Se define el **ángulo** formado entre dos planos como el ángulo formado por sus vectores normales.

Ejemplo 4. Encuentre el coseno del ángulo θ formado entre los planos

$$2x - y + z = 0,$$
$$x + 2y - z = 1.$$

Este coseno es el que corresponde al ángulo formado entre los vectores

$$A = (2, -1, 1) \quad \text{y} \quad B = (1, 2, -1).$$

Por consiguiente,

$$\cos\theta = \frac{A \cdot B}{\|A\| \, \|B\|} = -\frac{1}{6}.$$

Ejemplo 5. Sean

$$Q = (1, 1, 1) \qquad \text{y} \qquad P = (1, -1, 2).$$

Sea

$$N = (1, 2, 3).$$

Encuentre el punto de intersección de la recta que pasa por P en la dirección de N, con el plano que pasa por Q y que es perpendicular a N.

La representación paramétrica de la recta que pasa por P en la dirección de N es

(1) $$X = P + tN.$$

La ecuación del plano que pasa por Q y que es perpendicular a N es

(2) $$(X - Q) \cdot N = 0.$$

Podemos representar la recta y el plano de la manera siguiente:

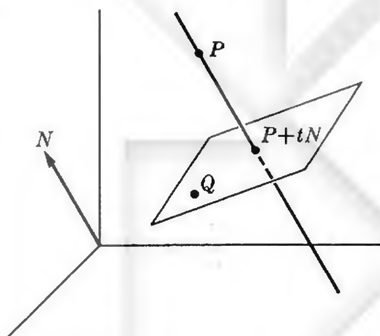


Figura 33

Debemos encontrar el valor de t tal que el vector X de (1) también satisfaga (2), esto es,

$$(P + tN - Q) \cdot N = 0,$$

o bien, después de usar las reglas del producto interior,

$$(P - Q) \cdot N + tN \cdot N = 0.$$

Al despejar t se obtiene

$$t = \frac{(Q - P) \cdot N}{N \cdot N} = \frac{1}{14},$$

por lo que el punto de intersección deseado es

$$P + tN = (1, -1, 2) + \tfrac{1}{14}(1, 2, 3) = (\tfrac{15}{14}, -\tfrac{12}{14}, \tfrac{31}{14}).$$

Ejemplo 6. Encuentre la ecuación del plano que pasa por los tres puntos siguientes:

$$P_1 = (1, 2, -1), \qquad P_2 = (-1, 1, 4), \qquad P_3 = (1, 3, -2).$$

Representamos en forma esquemática los tres puntos de la manera siguiente:

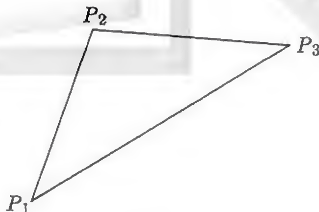


Figura 34

Luego encontramos un vector N perpendicular a $\overrightarrow{P_1P_2}$ y $\overrightarrow{P_1P_3}$ o, en otras palabras, perpendicular a $P_2 - P_1$ y $P_3 - P_1$. Tenemos que

$$P_2 - P_1 = (-2, -1, +5),$$

$$P_3 - P_1 = (0, 1, -1).$$

Sea $N = (a, b, c)$. Debemos resolver

$$N \cdot (P_2 - P_1) = 0 \qquad \text{y} \qquad N \cdot (P_3 - P_1) = 0,$$

en otras palabras,

$$-2a - b + 5c = 0,$$

$$b - c = 0.$$

Tomemos $b = c = 1$ y despejemos a; se obtiene $a = 2$. Entonces

$$N = (2, 1, 1)$$

satisface nuestros requerimientos. El plano perpendicular a N que pasa por P_1 es el plano deseado. Por consiguiente, su ecuación es $X \cdot N = P_1 \cdot N$, esto es,

$$2x + y + z = 2 + 2 - 1 = 3.$$

Distancia entre un punto y un plano. Considere un plano definido por la ecuación

$$(X - P) \cdot N = 0,$$

y sea Q un punto arbitrario. Queremos encontrar una fórmula para la distancia entre Q y el plano; con esto queremos decir la longitud del segmento desde Q hasta el punto de intersección de la recta perpendicular al plano que pasa por Q, tal como se muestra en la figura. Llamemos Q' a este punto de intersección.

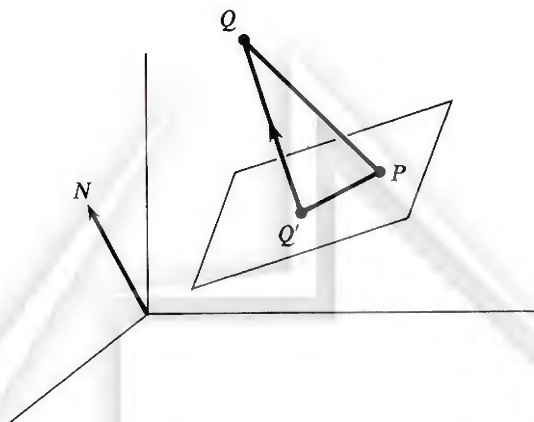


Figura 35

Por geometría, tenemos:

longitud del segmento $\overline{QQ'}$ = longitud de la proyección de $\overline{QP}$ sobre $\overline{QQ'}$.

Podemos expresar la longitud de esta proyección en términos del producto interior de la manera siguiente. Un vector unitario en la dirección de N, que es perpendicular al plano, está dado por $N/\|N\|$. Entonces

$$\begin{aligned}&\text{longitud de la proyección de } \overline{QP} \text{ sobre } \overline{QQ'}\\ &\qquad = \text{norma de la proyección de } Q-P \text{ sobre } N/\|N\|\\ &\qquad = \left|(Q-P)\cdot\frac{N}{\|N\|}\right|\end{aligned}$$

Ésta también se puede escribir en la forma:

$$\boxed{\text{distancia entre } Q \text{ y el plano} = \frac{|(Q-P)\cdot N|}{\|N\|}.}$$

Ejemplo 7. Sean

$$Q=(1,3,5), \qquad P=(-1,1,7) \qquad \text{y} \qquad N=(-1,1,-1).$$

La ecuación del plano es

$$-x+y-z=-5.$$

Encontramos que $\|N\|=\sqrt{3}$,

$$Q-P=(2,2,-2) \qquad \text{y} \qquad (Q-P)\cdot N=-2+2+2=2.$$

Por tanto, la distancia entre Q y el plano es igual a $2/\sqrt{3}$.

Ejercicios I, §6

1. Demuestre que las rectas $2x + 3y = 1$ y $5x - 5y = 7$ no son perpendiculares entre sí.

2. Sean $y = mx + b$ y $y = m'x + c$ las ecuaciones de dos rectas en el plano. Encuentre vectores perpendiculares a estas rectas. Demuestre que estos vectores son perpendiculares entre ellos si, y sólo si, $mm' = -1$.

Encuentre en el espacio de 2 dimensiones la ecuación de la recta perpendicular a N y que pasa por P, para los siguientes valores de N y P.

3. $N = (1, -1)$, $P = (-5, 3)$ 4. $= (-5, 4)$, $P = (3, 2)$

5. Demuestre que las rectas

$$3x - 5y = 1, \qquad 2x + 3y = 5$$

no son perpendiculares entre sí.

6. ¿Cuáles de las siguientes parejas de rectas son perpendiculares entre sí?
   (a) $3x - 5y = 1$ y $2x + y = 2$
   (b) $2x + 7y = 1$ y $x - y = 5$
   (c) $3x - 5y = 1$ y $5x + 3y = 7$
   (d) $-x + y = 2$ y $x + y = 9$

7. Encuentre la ecuación del plano perpendicular al vector N dado y que pasa por el punto P dado.
   (a) $N = (1, -1, 3)$, $P = (4, 2, -1)$
   (b) $N = (-3, -2, 4)$, $P = (2, \pi, -5)$
   (c) $N = (-1, 0, 5)$, $P = (2, 3, 7)$

8. Encuentre la ecuación del plano que pasa por los siguientes tres puntos.
   (a) $(2, 1, 1)$, $(3, -1, 1)$, $(4, 2, -1)$
   (b) $(-2, 3, -1)$, $(2, 2, 3)$, $(-4, -1, 1)$
   (c) $(-5, -1, 2)$, $(1, 2, -1)$, $(3, -1, 2)$

9. Encuentre un vector perpendicular a $(1, 2, -3)$ y $(2, -1, 3)$, y otro vector perpendicular a $(-1, 3, 2)$ y a $(2, 1, 1)$.

10. Encuentre un vector paralelo a la recta de intersección de los dos planos

$$2x - y + z = 1, \qquad 3x + y + z = 2.$$

11. La misma pregunta que en el ejercicio anterior, pero ahora para los planos

$$2x + y + 5z = 2, \qquad 3x - 2y + z = 3.$$

12. Encuentre una representación paramétrica para la recta de intersección de los planos de los ejercicios 10 y 11.

13. Encuentre el coseno del ángulo formado entre los siguientes planos:
   (a) $x + y + z = 1$, $x - y - z = 5$
   (b) $2x + 3y - z = 2$, $x - y + z = 1$
   (c) $x + 2y - z = 1$, $-x + 3y + z = 2$
   (d) $2x + y + z = 3$, $-x - y + z = \pi$

14. (a) Sean $P = (1, 3, 5)$ y $A = (-2, 1, 1)$. Encuentre la intersección de la recta que pasa por P en la dirección de A y el plano $2x + 3y - z = 1$.
(b) Sea $P = (1, 2, -1)$. Encuentre el punto de intersección del plano

$$3x - 4y + z = 2,$$

con la recta que pasa por P, perpendicular al plano.

15. Sean $Q = (1, -1, 2)$, $P = (1, 3, -2)$ y $N = (1, 2, 2)$. Encuentre el punto de intersección de la recta que pasa por P en la dirección de N y el plano que pasa por Q perpendicular a N.

16. Encuentre la distancia entre los puntos y los planos que se indican.
(a) $(1, 1, 2)$ y $3x + y - 5z = 2$
(b) $(-1, 3, 2)$ y $2x - 4y + z = 1$
(c) $(3, -2, 1)$ y el plano yz
(d) $(-3, -2, 1)$ y el plano yz

CAPÍTULO II

Matrices y ecuaciones lineales

El lector ya ha trabajado con ecuaciones lineales en sus cursos elementales. Las ecuaciones lineales simplemente son como las siguientes:

$$\begin{aligned} 2x + y + z &= 1, \\ 5x - y + 7z &= 0. \end{aligned}$$

El lector aprendió a resolver tales ecuaciones mediante la eliminación sucesiva de las variables. En este capítulo revisaremos la teoría de dichas ecuaciones, trabajando con ecuaciones en n variables e interpretando los resultados desde el punto de vista vectorial. Se darán algunas interpretaciones geométricas de las soluciones de las ecuaciones.

El primer capítulo se emplea muy poco en esta parte y se puede omitir por completo si tan sólo se conoce la definición de producto escalar entre dos n-tuplas. La multiplicación de matrices será formulada en términos de tal producto. Sin embargo, una interpretación geométrica de las soluciones de ecuaciones homogéneas se basará en el hecho de que el producto escalar entre dos vectores es igual a 0 si, y sólo si, los vectores son perpendiculares entre sí, de manera que, si el lector está interesado en esta interpretación, deberá consultar la sección donde se explica este hecho en el Capítulo I.

II, §1. Matrices

Consideremos una nueva clase de objetos, las matrices.

Sean n y m dos enteros ≥ 1. Un arreglo de números

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}$$

se conoce como **matriz**. Podemos abreviar la notación para esta matriz expresándola como (a_{ij}), $i = 1, \ldots, m$ y $j = 1, \ldots, n$. Decimos que es una matriz de m por n, o bien que es una matriz de $m \times n$. La matriz tiene m **renglones** y n **columnas**. Por ejemplo, la primera columna es

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$$

y el segundo renglón es $(a_{21}, a_{22}, \ldots, a_{2n})$. Decimos que a_{ij} es la **entrada** ij o la **componente** ij de la matriz.

Vea de nuevo el Capítulo I, §1. El ejemplo del espacio de 7 dimensiones tomado de la economía da lugar a una matriz (a_{ij}) de 7×7 $(i, j = 1, \ldots, 7)$, si definimos a_{ij} como la cantidad que la i-ésima industria gasta en la j-ésima industria. De modo que, manteniendo la notación de ese ejemplo, si $a_{25} = 50$, esto significa que la industria automotriz compró a la industria química 50 millones de dólares de materias primas durante un año determinado.

Ejemplo 1. La siguiente es una matriz de 2×3:

$$\begin{pmatrix} 1 & 1 & -2 \\ -1 & 4 & -5 \end{pmatrix}$$

Tiene dos renglones y tres columnas.

Los renglones son $(1, 1, -2)$ y $(-1, 4, -5)$. Las columnas son

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \quad \begin{pmatrix} -2 \\ -5 \end{pmatrix}$$

Por tanto, los renglones de una matriz se pueden considerar como n-tuplas y las columnas como m-tuplas verticales. Una m-tupla vertical también se conoce como **vector columna**.

Un vector $(x_1, \ldots, x_n)$ es una matriz de $1 \times n$. Un vector columna

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

es una matriz de $n \times 1$.

Cuando expresamos una matriz en la forma (a_{ij}), i denota el renglón y j la columna. De esta manera, en el ejemplo 1 tenemos

$$a_{11} = 1, \quad a_{23} = -5.$$

Un número individual (a) se puede considerar como una matriz de 1×1.

Sea (a_{ij}), $i = 1, \dots, m$ y $j = 1, \dots, n$, una matriz. Si $m = n$, entonces decimos que es una matriz **cuadrada**. Así,

$$\begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & -1 & 5 \\ 2 & 1 & -1 \\ 3 & 1 & -1 \end{pmatrix}$$

son matrices cuadradas.

Definimos la **matriz nula** como aquella en la que $a_{ij} = 0$ para todos i y j, y cuyo aspecto se muestra en seguida:

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

La representaremos con O. Observemos que hasta ahora nos hemos encontrado con el número cero, con el vector nulo y con la matriz nula.

Ahora vamos a definir la adición de matrices y la multiplicación de matrices por números.

Definimos la adición de matrices sólo cuando tienen el mismo tamaño. Así, sean m y n enteros fijos ≥ 1. Sean $A = (a_{ij})$ y $B = (b_{ij})$ matrices de $m \times n$. Definimos la matriz $A + B$ como aquella cuya componente en el renglón i y la columna j es $a_{ij} + b_{ij}$. En otras palabras, sumamos matrices del mismo tamaño, componente a componente.

Ejemplo 2. Sean

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 3 & 4 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 5 & 1 & -1 \\ 2 & 1 & -1 \end{pmatrix}$$

Entonces

$$A + B = \begin{pmatrix} 6 & 0 & -1 \\ 4 & 4 & 3 \end{pmatrix}$$

Si tanto A como B son matrices de $1 \times n$, esto es, n-tuplas, entonces observamos que nustra adición de matrices coincide con la adición que hemos definido en el Capítulo I para n-tuplas.

Si O es la matriz nula, entonces para cualquier matriz A (del mismo tamaño, por supuesto), tenemos que $O + A = A + O = A$.

Esto se puede verificar en forma trivial. Ahora definiremos la multiplicación de una matriz por un número. Sean c un número y $A = (a_{ij})$, una matriz. Definiremos cA como la matriz cuya componente ij es ca_{ij}. Escribimos

$$cA = (ca_{ij}).$$

Así pues, multiplicamos cada componente de A por c.

Ejemplo 3. Sean A y B como en el ejemplo 2. Sea $c = 2$. Entonces

$$2A = \begin{pmatrix} 2 & -2 & 0 \\ 4 & 6 & 8 \end{pmatrix} \quad \text{y} \quad 2B = \begin{pmatrix} 10 & 2 & -2 \\ 4 & 2 & -2 \end{pmatrix}.$$

Tenemos también que

$$(-1)A = -A = \begin{pmatrix} -1 & 1 & 0 \\ -2 & -3 & -4 \end{pmatrix}.$$

En general, para cualquier matriz $A = (a_{ij})$ representamos con $-A$ (menos A) la matriz $(-a_{ij})$. Como tenemos la relación $a_{ij} - a_{ij} = 0$ para números, también obtenemos la relación

$$A + (-A) = O$$

para matrices. También se conoce la matriz $-A$ como **inversa aditiva** de A.

Definimos otra noción relacionada con una matriz. Sea $A = (a_{ij})$ una matriz de $m \times n$. La matriz $B = (b_{ji})$ de $n \times m$ tal que $b_{ji} = a_{ij}$ se conoce como **transpuesta** de A, y se denota también con ${}^tA$. El transponer una matriz equivale a intercambiar renglones por columnas y viceversa. Si A es la matriz que aparece al principio de esta sección, entonces ${}^tA$ es la matriz

$$\begin{pmatrix} a_{11} & a_{21} & a_{31} & \cdots & a_{m1} \\ a_{12} & a_{22} & a_{32} & \cdots & a_{m2} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{mn} \end{pmatrix}.$$

Consideremos un caso particular:

$$\text{Si } A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 5 \end{pmatrix}, \quad \text{entonces} \quad {}^tA = \begin{pmatrix} 2 & 1 \\ 1 & 3 \\ 0 & 5 \end{pmatrix}.$$

Si $A = (2, 1, -4)$ es un *vector renglón*, entonces

$${}^tA = \begin{pmatrix} 2 \\ 1 \\ -4 \end{pmatrix}$$

es un *vector columna*.

Una matriz A que es igual a su transpuesta, esto es, $A = {}^tA$, se conoce como **simétrica**. Dicha matriz necesariamente es una matriz cuadrada.

Observación sobre la notación. He escrito el signo de transposición a la izquierda debido a que en muchas situaciones se considera la inversa de una matriz y se escribe A^{-1}, y entonces resulta más sencillo escribir ${}^tA^{-1}$ que $(A^{-1})^t$ o que $(A^t)^{-1}$ las que, de hecho, son iguales. No ha habido consenso en la comunidad matemática con respecto a dónde se debe colocar el signo de transposición, si a la derecha o a la izquierda.

Ejercicios II, §1

1. Sean

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 2 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} -1 & 5 & -2 \\ 1 & 1 & -1 \end{pmatrix}.$$

Encuentre $A+B$, $3B$, $-2B$, $A+2B$, $2A+B$, $A-B$, $A-2B$, $B-A$.

2. Sean

$$A = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} -1 & 1 \\ 0 & -3 \end{pmatrix}.$$

Encuentre $A+B$, $3B$, $-2B$, $A+2B$, $A-B$, $B-A$.

3. (a) Escriba los vectores renglón y los vectores columna de las matrices A y B del ejercicio 1.
   (b) Escriba los vectores renglón y los vectores columna de las matrices A y B del ejercicio 2.

4. (a) En el ejercicio 1, encuentre ${}^tA$ y ${}^tB$.
   (b) En el ejercicio 2, encuentre ${}^tA$ y ${}^tB$.

5. Si A y B son matrices arbitrarias de $m \times n$, demuestre que

$${}^t(A+B) = {}^tA + {}^tB.$$

6. Si c es un número, demuestre que ${}^t(cA) = c\,{}^tA$.

7. Si $A = (a_{ij})$ es una matriz cuadrada, entonces los elementos a_{ii} se denominan elementos diagonales. ¿En qué difieren los elementos diagonales de A y ${}^tA$?

8. En el ejercicio 2, determine ${}^t(A+B)$ y ${}^tA + {}^tB$.

9. En el ejercicio 2, determine $A + {}^tA$ y $B + {}^tB$.

10. (a) Demuestre que, para cualquier matriz cuadrada, la matriz $A + {}^tA$ es simétrica.
    (b) Se dice que una matriz A es antisimétrica si ${}^tA = -A$. Demuestre que, para cualquier matriz cuadrada A, la matriz $A - {}^tA$ es antisimétrica.
    (c) Si una matriz es antisimétrica, ¿qué puede usted decir acerca de sus elementos diagonales?

11. Sean

$$E_1 = (1,0,\dots,0), \quad E_2 = (0,1,0,\dots,0), \quad \dots, \quad E_n = (0,\dots,0,1)$$

los vectores unitarios canónicos de $\mathbf{R}^n$. Sean $x_1,\dots,x_n$ números. ¿Qué es $x_1E_1 + \cdots + x_nE_n$? Demuestre que, si

$$x_1E_1 + \cdots + x_nE_n = O,$$

entonces $x_i = 0$ para todo i.

II, §2. Multiplicación de matrices

Definiremos ahora el producto de matrices. Sea $A = (a_{ij})$, $i = 1, \ldots, m$ y $j = 1, \ldots, n$, una matriz de $m \times n$. Sea $B = (b_{jk})$, $j = 1, \ldots, n$ y sea $k = 1, \ldots, s$ una matriz de $n \times s$:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & \cdots & b_{1s} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{ns} \end{pmatrix}.$$

Definimos el **producto** AB como la matriz de $m \times s$ cuya coordenada ik es

$$\sum_{j=1}^{n} a_{ij} b_{jk} = a_{i1} b_{1k} + a_{i2} b_{2k} + \cdots + a_{in} b_{nk}.$$

Si $A_1, \ldots, A_m$ son los vectores renglón de la matriz A, y si $B^1, \ldots, B^s$ son los vectores columna de la matriz B, entonces la coordenada ik del producto AB es igual a $A_i \cdot B^k$. Así,

$$AB = \begin{pmatrix} A_1 \cdot B^1 & \cdots & A_1 \cdot B^s \\ \vdots & & \vdots \\ A_m \cdot B^1 & \cdot & A_m \cdot B^s \end{pmatrix}.$$

La multiplicación de matrices es, por consiguiente, una generalización del producto interior.

Ejemplo. Sean

$$A = \begin{pmatrix} 2 & 1 & 5 \\ 1 & 3 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & 4 \\ -1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Entonces AB es una matriz de 2×2 y los cálculos muestran que

$$AB = \begin{pmatrix} 2 & 1 & 5 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ -1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 15 & 15 \\ 4 & 12 \end{pmatrix}.$$

Ejemplo. Sea

$$C = \begin{pmatrix} 1 & 3 \\ -1 & -1 \end{pmatrix}.$$

Sean A y B como en el ejemplo 1. Entonces

$$BC = \begin{pmatrix} 3 & 4 \\ -1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 5 \\ -3 & -5 \\ 1 & 5 \end{pmatrix}$$

y

$$A(BC) = \begin{pmatrix} 2 & 1 & 5 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} -1 & 5 \\ -3 & -5 \\ 1 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 30 \\ -8 & 0 \end{pmatrix}.$$

Calcule $(AB)C$. ¿Qué encontró?

Si $X = (x_1, \ldots, x_m)$ es un vector renglón, es decir, una matriz de $1 \times m$, entonces podemos formar el producto XA, el cual tiene el siguiente aspecto:

$$(x_1, \ldots, x_m) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = (y_1, \ldots, y_n),$$

donde

$$y_k = x_1 a_{1k} + \cdots + x_m a_{mk}.$$

En este caso, XA es una matriz de $1 \times n$, es decir, un vector renglón.

Por otra parte, si X es un vector columna,

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

entonces $AX = Y$, donde Y también es un vector columna, cuyas coordenadas están dadas por

$$y_i = \sum_{j=1}^{n} a_{ij} x_j = a_{i1} x_1 + \cdots + a_{in} x_n.$$

La multiplicación $AX = Y$ tiene el siguiente aspecto:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Ejemplo. Ecuaciones lineales. Las matrices brindan una manera cómoda de escribir ecuaciones lineales. El lector seguramente ya ha trabajado con sistemas de ecuaciones lineales. Por ejemplo, una ecuación como:

$$3x - 2y + 3z = 1,$$

con tres incógnitas x, y, z. O bien un sistema de dos ecuaciones con tres incógnitas

(*)
$$\begin{aligned} 3x - 2y + 3z &= 1, \\ -x + 7y - 4z &= -5. \end{aligned}$$

En este ejemplo formamos la **matriz de los coeficientes**

$$A = \begin{pmatrix} 3 & -2 & 3 \\ -1 & 7 & -4 \end{pmatrix}.$$

Sea B el vector columna de los números que aparecen en el miembro derecho del sistema, es decir

$$B = \begin{pmatrix} 1 \\ -5 \end{pmatrix}.$$

Sea el vector de las incógnitas el siguiente vector columna:

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Entonces se percibe que el sistema de dos ecuaciones simultáneas se puede escribir en la siguiente forma:

$$AX = B.$$

Ejemplo. La primera ecuación de (*) representa la igualdad de las primeras componentes de AX y de B, mientras que la segunda ecuación de (*) representa la igualdad de las segundas componentes de AX y de B.

En general, sea $A = (a_{ij})$ una matriz de $m \times n$ y sea B un vector columna de tamaño m. Sea

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

un vector columna de tamaño n. Entonces el sistema de ecuaciones lineales

$$\begin{aligned} a_{11}x_1 &+ \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 &+ \cdots + a_{2n}x_n = b_2, \\ &\vdots \\ a_{m1}x_1 &+ \cdots + a_{mn}x_n = b_m, \end{aligned}$$

se puede escribir de la siguiente manera, que es más eficiente:

$$\boxed{AX = B,}$$

debido a la definición de multiplicación de matrices. Más adelante veremos cómo resolver tales sistemas. Decimos que hay m **ecuaciones** y n **incógnitas** o n **variables**.

Ejemplo. Matrices de Markov. A menudo se puede emplear una matriz para representar una situación práctica. Consideremos tres ciudades, digamos Los Ángeles, Chicago y Boston, que denotamos con LA, Ch y Bo. Supongamos que, en cualquier año dado, algunas personas salen de una de estas ciudades para ir a alguna de las otras. El porcentaje de las personas que salen y llegan está dado de la manera siguiente, por año:

$\frac{1}{4}$ LA va a Bo y $\frac{1}{7}$ LA va a Ch.

$\frac{1}{5}$ Ch va a LA y $\frac{1}{3}$ Ch va a Bo.

$\frac{1}{6}$ Bo va a LA y $\frac{1}{8}$ Bo va a Ch.

Sean x_n, y_n y z_n las poblaciones de LA, Ch y Bo, respectivamente, en el año n. Entonces podemos expresar la población en el año $n+1$ de la siguiente manera.

En el año $n+1$, $\frac{1}{4}$ de la población de LA sale para Boston y $\frac{1}{7}$ sale para Chicago. La fracción total que sale de LA durante el año es, por consiguiente,

$$\tfrac{1}{4} + \tfrac{1}{7} = \tfrac{11}{28}.$$

En consecuencia, la fracción total que permanece en LA es

$$1 - \tfrac{11}{28} = \tfrac{17}{28}.$$

Por tanto, la población de LA en el año $n+1$ es

$$x_{n+1} = \tfrac{17}{28}x_n + \tfrac{1}{5}y_n + \tfrac{1}{6}z_n.$$

En forma análoga, la fracción que sale de Chicago cada año es

$$\tfrac{1}{5} + \tfrac{1}{3} = \tfrac{8}{15},$$

de manera que la fracción que permanece en ese lugar es $\frac{7}{15}$. Por último, la fracción que sale de Boston cada año es

$$\tfrac{1}{6} + \tfrac{1}{8} = \tfrac{8}{24},$$

de manera que la fracción que permanece en Boston es $\frac{17}{24}$. Así,

$$\begin{aligned} y_{n+1} &= \tfrac{1}{7}x_n + \tfrac{7}{15}y_n + \tfrac{1}{8}z_n, \\ z_{n+1} &= \tfrac{1}{4}x_n + \tfrac{1}{3}y_n + \tfrac{17}{24}z_n. \end{aligned}$$

Sea A la matriz

$$A = \begin{pmatrix} \frac{17}{28} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{7} & \frac{7}{15} & \frac{1}{8} \\ \frac{1}{4} & \frac{1}{3} & \frac{17}{24} \end{pmatrix}.$$

Entonces podemos describir en forma más simple el cambio de población mediante la expresión

$$X_{n+1} = AX_n \qquad \text{donde} \qquad X_n = \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}.$$

El cambio de X_n a X_{n+1} se conoce como **proceso de Markov**. Éste se debe a la propiedad especial de la matriz A, cuyas componentes son todas ≥ 0 y tales que la suma de todos los elementos de cada columna es igual a 1. Una matriz así recibe el nombre de **matriz de Markov**.

Si A es una matriz cuadrada, entonces podemos formar el producto AA, que será una matriz cuadrada del mismo tamaño que A. Se denota con A^2. En forma análoga, podemos formar A^3, A^4 y, en general, A^n para cualquier entero positivo n. Por tanto, A^n es el producto de A consigo misma n veces.

Podemos definir la matriz unitaria de $n \times n$ como la matriz que tiene todas las componentes diagonales iguales a 1 y todas las demás componentes iguales a

0. Así, la matriz unitaria de $n \times n$, denotada con I_n, tiene el siguiente aspecto:

$$I_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

Entonces podemos definir $A^0 = I$ (la matriz unitaria del mismo tamaño que A). Observe que, para cualesquiera dos enteros r y $s \geq 0$, tenemos la relación usual

$$A^r A^s = A^s A^r = A^{r+s}.$$

Por ejemplo, en el proceso de Markov descrito anteriormente, podemos expresar el vector de población en el año $n+1$ de la siguiente manera:

$$X_{n+1} = a^n X_1,$$

donde X_1 es el vector de población en el primer año.

Advertencia. No siempre es cierto que $AB = BA$. Por ejemplo, calcule AB y BA en el siguiente caso:

$$A = \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & -1 \\ 0 & 5 \end{pmatrix}.$$

Encontrará dos valores diferentes. Esto se expresa diciendo que la multiplicación de matrices no necesariamente es conmutativa. Desde luego, en algunos casos *especiales*, sí tenemos que $AB = BA$. Por ejemplo, las potencias de A conmutan, es decir, tenemos que $A^r A^s = A^s A^r$, tal como se indicó antes.

Probemos ahora otras propiedades básicas de la multiplicación.

Ley distributiva. *Sean A, B y C matrices. Supongamos que A y B se pueden multiplicar entre sí, que A y C se pueden multiplicar entre sí y que B y C se pueden sumar. Entonces A y $B+C$ se pueden multiplicar entre sí, y tenemos que*

$$A(B+C) = AB + AC.$$

Si x es un número, entonces

$$A(xB) = x(AB).$$

Demostración. Sea A_i el i-ésimo renglón de A y sean B^k y C^k las k-ésimas columnas de B y C, respectivamente.... Entonces k-ésima columna de $B+C$ es $B^k + C^k$. Por definición, la componente ik de $A(B+C)$ es $A_i \cdot (B^k + C^k)$. Como

$$A_i \cdot (B^k + C^k) = A_i \cdot B^k + A_i \cdot C^k,$$

se infiere nuestra primera afirmación. Con respecto a la segunda, observe que la k-ésima columna de xB es xB^k. Como

$$A_i \cdot xB^k = x(A_i \cdot B^k),$$

se infiere nuestra segunda afirmación.

Ley asociativa. *Sean A, B y C matrices tales que A y B se pueden multiplicar entre sí y que B y C se pueden multiplicar entre sí. Entonces A y BC se pueden multiplicar entre sí. Lo mismo sucede con AB y C y tenemos*

$$(AB)C = A(BC).$$

Demostración. Sea $A = (a_{ij})$ una matriz de $m \times n$; sea $B = (b_{jk})$ una matriz de $n \times r$ y sea $C = (c_{kl})$ una matriz de $r \times s$. El producto AB es una matriz de $m \times r$, cuya componente ik es igual a la suma

$$a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk}.$$

Abreviaremos esta suma mediante el empleo de nuestra notación $\sum$ escribiendo

$$\sum_{j=1}^{n} a_{ij}b_{jk}.$$

Por definición, la componente il de $(AB)C$ es igual a

$$\sum_{k=1}^{r}\left[\sum_{j=1}^{n} a_{ij}b_{jk}\right]c_{kl} = \sum_{k=1}^{r}\left[\sum_{j=1}^{n} a_{ij}b_{jk}c_{kl}\right]$$

La suma que aparece a la derecha de la igualdad tambien se puede describir como la suma de todos los términos

$$\sum a_{ij}b_{jk}c_{kl},$$

donde j y k varían sobre todos los enteros $1 \leq j \leq n$ y $1 \leq k \leq r$, respectivamente.

Si hubiéramos comenzado con la componente jl de BC y luego hubiéramos calculado la componente il de $A(BC)$, habríamos encontrado exactamente la misma suma, probando de esa manera la propiedad deseada.

Las propiedades anteriores son muy similares a las de la multiplicación de números, excepto que no se cumple la ley conmutativa.

También podemos relacionar la multiplicación con la transpuesta:

Sean A y B matrices de un tamaño tal que AB está definida. Entonces

$${}^t(AB) = {}^tB\,{}^tA.$$

En otras palabras, la transpuesta del producto es igual al producto de las transpuestas en orden inverso.

Demostración. Sea $A = (a_{ij})$ y $B = (b_{jk})$. Entonces $AB = C = (c_{ik})$, donde

$$\begin{aligned} c_{ik} &= a_{i1}b_{1k} + \cdots + a_{in}b_{nk} \\ &= b_{1k}a_{i1} + \cdots + b_{nk}a_{in}. \end{aligned}$$

Sean ${}^tA = (a'_{ji})$, ${}^tB = (b'_{kj})$ y ${}^tC = (c'_{ki})$. Entonces

$$a'_{ji} = a_{ij}, \qquad b'_{kj} = b_{jk}, \qquad c'_{ki} = c_{ik}.$$

En consecuencia, podemos escribir la relación anterior como sigue:

$$c'_{ki} = b'_{kl}a'_{li} + \cdots + b'_{kn}a'_{ni},$$

lo que muestra que ${}^tC = {}^tB\,{}^tA$, tal como se deseaba.

Ejemplo. En lugar de escribir el sistema de ecuaciones lineales $AX = B$ en términos de vectores columna, podemos escribirlo considerando las transpuestas, lo que da por resultado

$${}^tX\,{}^tA = {}^tB.$$

Si X y B son vectores columna, entonces ${}^tX$ y ${}^tB$ son vectores renglón. A veces conviene reescribir el sistema de esta manera.

A diferencia de la división entre números no nulos, **no podemos dividir entre una matriz**, como tampoco podemos dividir entre un vector (n-tupla). En ciertas circunstancias, podemos definir una inversa de la siguiente manera. Esto lo hacemos sólo para matrices cuadradas. Sea A una matriz de $n \times n$. Una **inversa de** A es una matriz B tal que

$$AB = BA = I.$$

Puesto que multiplicamos A por B en ambos lados, la única manera de que esto tenga sentido es que B también sea una matriz de $n \times n$. Algunas matrices no tienen inversas. Sin embargo, **si existe una inversa, entonces existe sólo una** (decimos que **la inversa es única, o que está determinada en forma única por** A). Esto es fácil de probar. Supongamos que B y C son inversas de A, de manera que

$$AB = BA = I \quad \text{y} \quad AC = CA = I.$$

Multipliquemos la ecuación $BA = I$ a la derecha por C. Entonces

$$BAC = IC = C$$

y hemos supuesto que $AC = I$, de manera que $BAC = BI = B$. Esto prueba que $B = C$. A la luz de lo anterior, denotamos la inversa con

$$A^{-1}.$$

Entonces A^{-1} es la única matriz que

$$\boxed{A^{-1}A = I \quad \text{y} \quad AA^{-1} = I.}$$

Más adelante probaremos que, si A y B son matrices cuadradas del mismo tamaño tal que $AB = I$, entonces se infiere que también

$$BA = I.$$

En otras palabras, si B es una inversa por la derecha de A, entonces también es una inversa por la izquierda. De momento, el lector puede suponer esto. Así, cuando verifique que una matriz es la inversa de otra, sólo necesita hacerlo por un lado.

También encontraremos más adelante una manera de calcular la inversa, cuando exista, lo que puede ser un asunto tedioso.

Sea c un número. Entonces la matriz

$$cI = \begin{pmatrix} c & 0 & \cdots\cdots & 0 \\ 0 & c & 0 \ \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots\cdots & \cdots & c \end{pmatrix}$$

que tiene componente c en cada entrada diagonal y 0 en las demás entradas, se conoce como **matriz escalar**. También la podemos escribir como cI, donde I es la matriz unitaria de $n \times n$. Véase el ejercicio 6.

Como una aplicación de la fórmula para la transpuesta de un producto, veremos ahora que:

La transpuesta de una inversa es la inversa de la transpuesta, esto es

$${}^t(A^{-1}) = ({}^tA)^{-1}.$$

Demostración. Consideremos la transpuesta de la relación $AA^{-1} = I$. Entonces, por la regla para la transpuesta de un producto, obtenemos

$${}^t(A^{-1}){}^tA = {}^tI = I$$

porque I es igual a su propia transpuesta. Del mismo modo, al aplicar la transpuesta a la relación $A^{-1}A = I$ se obtiene

$${}^tA{}^t(A^{-1}) = {}^tI = I.$$

Por tanto, ${}^t(A^{-1})$ es una inversa de ${}^tA$, como se quería mostrar.

En vista de este resultado, se acostumbra omitir los paréntesis y se escribe

$${}^tA^{-1}$$

para representar la inversa de la transpuesta, la que, según hemos visto, es igual a la transpuesta de la inversa.

Finalizamos esta sección con un ejemplo importante de multiplicación de matrices.

Ejemplo. Rotaciones. Un tipo especial de matrices de 2×2 representa rotaciones. Para cada número θ, sea $R(\theta)$ la matriz

$$R(\theta) = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

Sea $X = \begin{pmatrix} x \\ y \end{pmatrix}$ un punto sobre el círculo unitario. Podemos escribir sus coordenadas x, y en la forma

$$x = \cos\varphi, \qquad y = \operatorname{sen}\varphi$$

para algún número φ. Entonces obtenemos, mediante la multiplicación de matrices,

$$R(\theta)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \cos\varphi \\ \operatorname{sen}\varphi \end{pmatrix}$$

$$= \begin{pmatrix} \cos(\theta+\varphi) \\ \operatorname{sen}(\theta+\varphi) \end{pmatrix}.$$

Esto se infiere de las **fórmulas de adición para el seno y el coseno**, a saber,

$$\cos(\theta + \varphi) = \cos\theta\cos\varphi - \operatorname{sen}\theta\operatorname{sen}\varphi,$$
$$\operatorname{sen}(\theta + \varphi) = \operatorname{sen}\theta\cos\varphi + \cos\theta\operatorname{sen}\varphi.$$

Un punto arbitrario de $\mathbf{R}^2$ se puede escribir en la forma

$$rX = \begin{pmatrix} r\cos\varphi \\ r\operatorname{sen}\varphi \end{pmatrix},$$

donde r es un número ≥ 0. Como

$$R(\theta)rX = rR(\theta)X,$$

vemos que la multiplicación por $R(\theta)$ también tiene el efecto de hacer girar a rX en un ángulo θ. Así, la rotación en un ángulo θ se puede representar por la matriz $R(\theta)$.

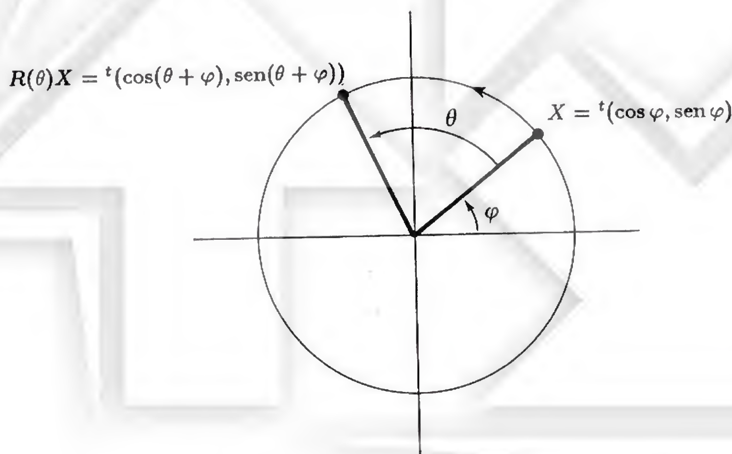


Figura 1

Observe que, por razones tipográficas, hemos escrito el vector ${}^tX$ en forma horizontal, aunque hemos puesto una pequeña t en el superíndice superior izquierdo, para denotar la transpuesta, de manera que X es un vector columna.

Ejemplo. La matriz que corresponde a la rotación en un ángulo de $\pi/3$ está dada por

$$R(\pi/3) = \begin{pmatrix} \cos\pi/3 & -\operatorname{sen}\pi/3 \\ \operatorname{sen}\pi/3 & \cos\pi/3 \end{pmatrix}$$
$$= \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}.$$

Ejemplo S. ea $X = {}^t(2,5)$. Giremos a X en un ángulo de $\pi/3$, y encontremos las coordenadas del vector girado.

Estas coordenadas son:

$$R(\pi/3)X = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix} \begin{pmatrix} 2 \\ 5 \end{pmatrix}$$
$$= \begin{pmatrix} 1 - 5\sqrt{3}/2 \\ \sqrt{3} + 5/2 \end{pmatrix}.$$

Advertencia. Observe cómo multiplicamos por la izquierda al vector columna por la matriz $R(\theta)$. Si el lector quiere trabajar con vectores renglón, entonces considere la transpuesta y verifique directamente que

$$(2,5)\begin{pmatrix} 1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{pmatrix} = (1 - 5\sqrt{3}/2, \sqrt{3} + 5/2).$$

En este caso se ha transpuesto la matriz $R(\theta)$. El signo menos ahora aparece en la esquina inferior izquierda.

Ejercicios II, §2

Los siguientes ejercicios ayudan principalmente a mecanizar la multiplicación de matrices. Sin embargo, también ilustra algunos aspectos más teóricos de esta multiplicación. Por consiguiente, deberán resolverse todos. En forma más específica, tenemos que:

Del ejercicio 7 al 12 se ejemplifica la multiplicación por los vectores unitarios canónicos.

Del ejercicio 14 al 19 se ejemplifica la multiplicación de matrices triangulares.

Del ejercicio 24 al 27 se ejemplifica la transformación de la adición de números en multiplicación de matrices.

Del ejercicio 27 al 32 se ejemplifican las rotaciones.

Del ejercicio 33 al 37 se ejemplifican las matrices elementales *y deberán resolverse antes de estudiar la sección* §5.

1. Sea I la matriz unitaria de $n \times n$. Sea A una matriz de $n \times r$. ¿A qué es igual IA? Si A es una matriz de $m \times n$, ¿a qué es igual AI?

2. Sea O la matriz cuyas coordenadas son todas iguales a 0. Sea A una matriz de un tamaño tal que esté definido el producto AO. ¿A qué es igual AO?

3. En cada uno de los siguientes casos, encuentre $(AB)C$ y $A(BC)$.

   (a) $A = \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$

   (b) $A = \begin{pmatrix} 2 & 1 & -1 \\ 3 & 1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 3 & -1 \end{pmatrix}$, $C = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$

   (c) $A = \begin{pmatrix} 2 & 4 & 1 \\ 3 & 0 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & -1 \\ 3 & 1 & 5 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 2 \\ 3 & 1 \\ -1 & 4 \end{pmatrix}$

4. Sean A y B matrices cuadradas del mismo tamaño y suponga que $AB = BA$. Demuestre que

$$(A+B)^2 = A^2 + 2AB + B^2, \qquad \text{y} \qquad (A+B)(A-B) = A^2 - B^2,$$

usando la ley distributiva.

5. Sean

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}.$$

Encuentre AB y BA.

6. Sea

$$C = \begin{pmatrix} 7 & 0 \\ 0 & 7 \end{pmatrix}.$$

Sean A y B como en el ejercicio 5. Encuentre CA, AC, CB y BC. Establezca la regla general incluyendo este ejercicio como caso especial.

7. Sean $X = (1, 0, 0)$ y

$$A = \begin{pmatrix} 3 & 1 & 5 \\ 2 & 0 & 1 \\ 1 & 1 & 7 \end{pmatrix}.$$

¿Qué es XA?

8. Sea $X = (0, 1, 0)$ y sea A una matriz arbitraria de 3×3. ¿De qué manera se podría describir XA? ¿Y si $X = (0, 0, 1)$? Generalice obteniendo enunciados similares referidos a matrices de $n \times n$ y a sus productos con vectores unitarios.

9. Sea

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 1 & 5 \end{pmatrix}.$$

Encuentre AX para cada uno de los siguientes valores de X.

(a) $X = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ (b) $X = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ (c) $X = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

10. Sea

$$A = \begin{pmatrix} 3 & 7 & 5 \\ 1 & -1 & 4 \\ 2 & 1 & 8 \end{pmatrix}.$$

Encuentre AX para cada uno de los valores de X dados en el ejercicio 9.

11. Sean

$$X = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad \text{y} \qquad A = \begin{pmatrix} a_{11} & \cdots & a_{14} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{m4} \end{pmatrix}.$$

¿Qué es AX?

12. Sea X un vector columna con todas sus componentes iguales a 0 excepto la componente j, que es igual a 1. Sea A una matriz arbitraria, cuyo tamaño es tal que podemos formar el producto AX. ¿A qué es igual AX?

13. Sean X el vector columna y A la matriz indicados en cada caso. Encuentre AX como vector columna.

(a) $X = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$, $A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 2 & 0 & -1 \end{pmatrix}$ (b) $X = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $A = \begin{pmatrix} 2 & 1 & 5 \\ 0 & 1 & 1 \end{pmatrix}$

(c) $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ (d) $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$

14. Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Encuentre el producto AS para cada una de las siguientes matrices S. Describa con palabras el efecto sobre A de este producto.

(a) $S = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$ (b) $S = \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}$.

15. De nuevo, sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Encuentre el producto SA para cada una de las siguientes matrices S. Describa con palabras el efecto de este producto sobre A.

(a) $S = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$ (b) $S = \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}$.

16. (a) Sea A la matriz

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Encuentre A^2 y A^3. Generalice a matrices de 4×4.

(b) Sea A la matriz

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Calcule A^2, A^3 y A^4.

17. Sea

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Encuentre A^2, A^3 y A^4.

18. Sea A una matriz diagonal, cuyos elementos diagonales son $a_1, \ldots, a_n$. ¿A qué son iguales A^2, A^3 y A^k para cualquier entero positivo k?

19. Sea

$$A = \begin{pmatrix} 0 & 1 & 6 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}.$$

Encuentre A^3.

20. (a) Encuentre una matriz A de 2×2 tal que $A^2 = -I = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$.

(b) Determine todas las matrices A de 2×2 tales que $A^2 = O$.

21. Sea A una matriz cuadrada.
(a) Si $A^2 = O$, demuestre que $I - A$ es invertible.
(b) Si $A^3 = O$, demuestre que $I - A$ es invertible.
(c) En general, si $A^n = O$ para algún entero positivo n, demuestre que $I - A$ es invertible. (*Sugerencia*: Considere la serie geométrica.)
(d) Suponga que $A^2 + 2A + I = O$. Demuestre que A es invertible.
(e) Suponga que $A^3 - A + I = O$. Demuestre que A es invertible.

22. Sean A y B dos matrices cuadradas del mismo tamaño. Decimos que A es **semejante** a B si existe una matriz invertible T tal que $B = TAT^{-1}$. Suponga que éste es el caso. Pruebe que:
(a) B es semejante a A.
(b) A es invertible si, y sólo si, B es invertible.
(c) ${}^tA$ es semejante a ${}^tB$.
(d) Suponga que $A^n = O$ y que B es una matriz invertible del mismo tamaño que A. Demuestre que $(BAB^{-1})^n = O$.

23. Sea A una matriz cuadrada de la forma

$$\begin{pmatrix} a_{11} & * & \cdots\cdots & * \\ 0 & a_{22} & * \ \cdots & * \\ \vdots & & \ddots & \vdots \\ & & & * \\ 0 & \cdots\cdots & 0 & a_{nn} \end{pmatrix}.$$

La notación significa que todos los elementos que se encuentran por debajo de la diagonal son iguales a 0 y que los elementos que se encuentran arriba de la diagonal son arbitrarios. Se puede expresar esta propiedad diciendo que

$$a_{ij} = 0 \quad \text{si} \quad i > j.$$

Una matriz así se conoce como **superiomente triangular**. Si A y B son matrices superiormente triangulares (del mismo tamaño), ¿qué puede decirse de los elementos diagonales de AB?

En los ejercicios 24 a 27 se dan ejemplos en los que la adición de números se transforma en multiplicación de matrices.

24. Sean a y b números, y sean

$$A = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}.$$

¿A qué es igual AB? ¿A qué son iguales A^2 y A^3? ¿Qué es A^n, donde n es un entero positivo?

25. Demuestre que la matriz A del ejercicio 24 tiene una inversa. ¿Cuál es esta inversa?

26. Demuestre que, si A y B son matrices de $n \times n$ que tienen inversas, entonces AB tiene una inversa.

27. **Rotaciones**. Sea $R(\theta)$ la matriz

$$R(\theta) = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

(a) Demuestre que, para cualesquiera dos números θ_1 y θ_2, tenemos que

$$R(\theta_1)R(\theta_2) = R(\theta_1 + \theta_2).$$

[El lector tendrá que usar las fórmulas de adición para el seno y el coseno.]

(b) Demuestre que la matriz $R(\theta)$ tiene una inversa y descríbala.

(c) Sea $A = R(\theta)$. Demuestre que

$$A^2 = \begin{pmatrix} \cos 2\theta & -\operatorname{sen} 2\theta \\ \operatorname{sen} 2\theta & \cos 2\theta \end{pmatrix}.$$

(d) Determine A^n para cualquier entero positivo n. Use inducción.

28. Encuentre la matriz $A = R(\theta)$ asociada con la rotación para cada uno de los siguientes valores de θ.

(a) $\pi/2$ (b) $\pi/4$ (c) π (d) $-\pi$ (e) $-\pi/3$
(f) $\pi/6$ (g) $5\pi/4$

29. En general, sea $\theta > 0$. ¿Cuál es la matriz asociada con la rotación en un ángulo $-\theta$ (es decir, una rotación en un ángulo θ en sentido dextrógiro)?

30. Sea $X = {}^t(1,2)$ un punto del plano. Si usted gira a X en un ángulo de $\pi/4$, ¿cuáles son las coordenadas del nuevo punto?

31. La misma pregunta que en el ejercicio anterior, pero ahora cuando $X = {}^t(-1,3)$ y se gira en un ángulo de $\pi/2$.

32. Para cualquier vector X de $\mathbf{R}^2$, sea $Y = R(\theta)X$ su rotación en un ángulo θ. Demuestre que $\|Y\| = \|X\|$.

Los siguientes ejercicios sobre matrices elementales deberán realizarse antes de estudiar la sección §5.

33. **Matrices elementales**. Sea

$$A = \begin{pmatrix} 2 & 3 & -1 & 1 \\ 1 & 4 & 2 & -2 \\ -1 & 1 & 3 & -5 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

Sea U la matriz que se muestra en cada inciso. En cada caso encuentre UA.

(a) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

(c) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ (d) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

(e) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ (f) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

34. Sea E la matriz que se muestra en cada inciso. Encuentre EA, donde A es la misma matriz que se mencionó en el ejercicio anterior.

(a) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 5 & 0 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

35. Sea E la matriz mostrada en cada inciso. Encuentre EA, donde A es la misma matriz que se mencionó en el ejercicio anterior y en el ejercicio 33.

(a) $\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

36. Sea $A = (a_{ij})$ una matriz de $m \times n$,

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

Sean $1 \leq r \leq m$ y $1 \leq s \leq m$. Sea I_{rs} la matriz cuya componente rs es igual a 1 y tal que todas las otras componentes son iguales a 0.

(a) ¿A qué es igual $I_{rs}A$?
(b) Suponga que $r \neq s$. ¿A qué es igual $(I_{rs} + I_{sr})A$?
(c) Suponga que $r \neq s$. Sea I_{jj} la matriz cuya componete jj es igual a 1 y tal que las otras componentes son iguales a 0. Sea

$$E_{rs} = I_{rs} + I_{sr} + \text{ suma de todas las } I_{jj} \text{ para } j \neq r,\ j \neq s.$$

¿A qué es igual $E_{rs}A$?

37. De nuevo, sea $r \neq s$.
(a) Sea $E = I + 3I_{rs}$. ¿A qué es igual EA?
(b) Sea c cualquier número. Sea $E = I + cI_{rs}$. ¿A qué es igual EA?

El resto del capítulo estará relacionado principalmente con ecuaciones lineales y, en especial, con las homogéneas. Encontraremos tres maneras de interpretar tales ecuaciones, lo que brindará tres maneras diferentes de considerar las matrices y los vectores.

II, §3. Ecuaciones lineales homogéneas y eliminación

En esta sección consideraremos las ecuaciones lineales mediante un método de eliminación. En la siguiente, analizaremos otro método.

Nos interesará el caso en que el número de incógnitas es mayor que el número de ecuaciones y veremos que, en ese caso, siempre existe una solución no trivial.

Antes de tratar el caso general, estudiaremos ejemplos.

Ejemplo 1. Suponga que tenemos una sola ecuación, como la siguiente:

$$2x + y - 4z = 0.$$

Deseamos encontrar una solución tal que x, y y z no sean simultáneamente iguales a cero. Una ecuación equivalente es

$$2x = -y + 4z.$$

Para encontrar una solución no trivial, damos a todas las variables, excepto la primera, un valor especial $\neq 0$, digamos $y = 1$ y $z = 1$. Entonces despejamos x. Encontramos que

$$2x = -y + 4z = 3,$$

por lo que $x = \frac{2}{3}$.

Ejemplo 2. Consideremos una pareja de ecuaciones, digamos

(1) $$2x + 3y - z = 0,$$

(2) $$x + y + z = 0.$$

Mediante la eliminación de una variable, reducimos el problema de resolver estas ecuaciones simultáneas al caso anterior de una ecuación. Así, multiplicamos la segunda ecuación por 2 y la restamos de la primera ecuación, con lo que obtenemos

(3) $$y - 3z = 0.$$

Ahora nos encontramos con una ecuación con más de una variable. Damos a z cualquier valor $\neq 0$, digamos $z = 1$, y despejamos y, a saber, $y = 3$. Luego despejamos x a partir de la segunda ecuación, a saber, $x = -y - z$ y obtenemos $x = -4$. Los valores que hemos obtenido para x, y y z también son soluciones de la primera ecuación, debido a que la primera ecuación es (en un sentido obvio) la suma de la ecuación (2) multiplicada por 2 y la ecuación (3).

Ejemplo 3. Deseamos encontrar una solución para el sistema de ecuaciones

$$3x - 2y + z + 2w = 0,$$

$$x + y - z - w = 0,$$

$$2x - 2y + 3z = 0.$$

De nuevo empleamos el método de eliminación. Multiplicamos la segunda ecuación por 2 y la restamos de la tercera. Encontramos que

$$-4y + 5z + 2w = 0.$$

Multiplicamos la segunda ecuación por 3 y la restamos de la primera. Encontramos que

$$-5y + 4z + 5w = 0.$$

Ahora hemos eliminado x de nuestras ecuaciones y encontramos dos ecuaciones con tres incógnitas, y, z y w. Eliminamos y de estas dos ecuaciones de la manera siguiente: multiplicamos la superior por 5, multiplicamos la inferior por 4 y las restamos. Obtenemos

$$9z - 10w = 0.$$

Ahora demos a w un valor arbitrario $\neq 0$, digamos $w = 1$. Entonces podemos despejar z, a saber,

$$z = 10/9.$$

Regresando a las ecuaciones inmediatamente anteriores, despejemos y usando

$$4y = 5z + 2w.$$

Esto da como resultado

$$y = 68/9.$$

Por último, despejemos x empleando, por ejemplo, la segunda ecuación del conjunto original de tres, de manera que

$$x = -y + z + w,$$

o, en forma numérica,

$$x = -49/9.$$

De modo que hemos encontrado:

$$w = 1, \qquad z = 10/9, \qquad y = 68/9, \qquad x = -49/9.$$

Observemos que teníamos tres ecuaciones con cuatro incógnitas. Mediante una eliminación sucesiva de variables, redujimos estas ecuaciones a dos ecuaciones con tres incógnitas y luego a una ecuación con dos incógnitas.

Usando precisamente el mismo método, suponga que comenzamos con tres ecuaciones con cinco incógnitas. Al eliminar una variable se obtendrán dos ecuaciones con cuatro incógnitas. La eliminación de otra variable dará por resultado una ecuación con tres incógnitas. Entonces podemos resolver esta ecuación y volver hacia atrás para obtener valores para las variables previas, tal como hemos mostrado en los ejemplos.

En general, suponga que comenzamos con m ecuaciones con n incógnitas y $n > m$. Eliminamos una de las variables, digamos x_1, y obtenemos un sistema de $m - 1$ ecuaciones con $n - 1$ incógnitas. Eliminamos una segunda variable,

digamos x_2, y obtenemos un sistema de $m-2$ ecuaciones con $n-2$ incógnitas. Mediante la repetición de este proceso eliminamos $m-1$ variables, y terminamos con 1 ecuación con $n-m+1$ incógnitas. Entonces damos valores arbitrarios no triviales a todas las variables restantes excepto una, despejamos esta última variable y luego procedemos hacia atrás a fin de despejar en forma sucesiva cada una de las variables eliminadas, como hicimos en nuestros ejemplos. De este modo hemos obtenido una manera eficaz para encontrar una solución no trivial del sistema original.

Podemos decir todo esto de manera precisa en términos de inducción.

Sea $A = (a_{ij})$, $i = 1, \ldots, m$ y $j = 1, \ldots, n$, una matriz. Sean $b_1, \ldots, b_m$ números. Ecuaciones como

$$(*) \qquad \begin{matrix} a_{11}x_1 & + \cdots + & a_{1n}x_n & = b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1}x_1 & + \cdots + & a_{mn}x_n & = b_m \end{matrix}$$

se conocen como ecuaciones lineales. También decimos que $(*)$ es un sistema de ecuaciones lineales. Se dice que el sistema es **homogéneo** si todos los números $b_1, \ldots, b_m$ son iguales a 0. El número n se conoce como el número de **incógnitas** y m es el número de ecuaciones.

El sistema de ecuaciones

$$(**) \qquad \begin{matrix} a_{11}x_1 & + \cdots + & a_{1n}x_n & = 0 \\ \vdots & & \vdots & \\ a_{m1}x_1 & + \cdots + & a_{mn}x_n & = 0 \end{matrix}$$

se llamará **sistema homogéneo asociado con** $(*)$. En esta sección estudiaremos el sistema homogéneo $(**)$.

El sistema $(**)$ siempre tiene una solución, a saber, la solución obtenida al hacer todo $x_i = 0$. Esta solución se denominará **solución trivial**. Una solución $(x_1, \ldots, x_n)$ tal que algún x_i es $\neq 0$ se conoce como **no trivial**.

Consideremos nuestro sistema de ecuaciones homogéneas $(**)$. Sean $A_1, \ldots, A_m$ los vectores renglón de la matriz (a_{ij}). Entonces podemos escribir nuestras ecuaciones $(**)$ en la siguiente forma:

$$(**) \qquad \begin{matrix} A_1 \cdot X = 0 \\ \vdots \\ A_m \cdot X = 0. \end{matrix}$$

Por consiguiente, una solución del sistema de ecuaciones lineales se puede interpretar como el conjunto de todas las n-tuplas X que son perpendiculares a los vectores renglón de la matriz A. Geométricamente, encontrar una solución de $(**)$ equivale a encontrar un vector que es perpendicular a $A_1, \ldots, A_m$. Al usar la notación del producto interior se hará más fácil formular la demostración de nuestro teorema principal, a saber:

Teorema 3.1. *Sea*

$$(**) \qquad \begin{matrix} a_{11}x_1 & + \cdots + & a_{1n}x_n & = 0 \\ \vdots & & \vdots & \\ a_{m1}x_1 & + \cdots + & a_{mn}x_n & = 0 \end{matrix}$$

un sistema de m ecuaciones lineales con n incógnitas y supongamos que $n > m$. Entonces el sistema tiene una solución no trivial.

Demostración. La demostración se llevará a cabo por inducción.

Primero consideremos el caso de una ecuación con n incógnitas, $n > 1$;

$$a_1x_1 + \cdots + a_nx_n = 0.$$

Si todos los coeficientes $a_1, \ldots, a_n$ son iguales a 0, entonces cualquier valor de las variables será una solución y, ciertamente, existe una solución no trivial. Supongamos que algún coeficiente a_i es $\neq 0$. Después de reenumerar las variables y los coeficientes, podemos suponer que este coeficiente es a_1. Entonces damos a $x_2, \ldots, x_n$ valores arbitrarios, por ejemplo, $x_2 = \cdots = x_n = 1$, y despejamos x_1, obteniendo

$$x_1 = \frac{-1}{a_1}(a_2 + \cdots + a_n).$$

De esta manera obtenemos una solución no trivial para nuestro sistema de ecuaciones.

Ahora supongamos que nuestro teorema es cierto para un sistema de $m - 1$ ecuaciones con más de $m - 1$ incógnitas. Probaremos que es cierto para m ecuaciones con n incógnitas cuando $n > m$. Consideremos el sistema (**).

Si todos los coeficientes (a_{ij}) son iguales a 0, podemos dar a nuestras variables cualquier valor no nulo con el objeto de obtener una solución. Si algún coeficiente no es igual a 0, entonces, luego de reenumerar las ecuaciones y las variables, podemos suponer que es a_{11}. Para eliminar x_1 restamos un múltiplo de la primera ecuación de las otras ecuaciones. A saber, consideramos el sistema de ecuaciones

$$\begin{aligned} \left(A_2 - \frac{a_{21}}{a_{11}}A_1\right) \cdot X &= 0 \\ &\vdots \\ \left(A_2 - \frac{a_{m1}}{a_{11}}A_1\right) \cdot X &= 0, \end{aligned}$$

que también se puede escribir en la forma

$$(***) \qquad \begin{aligned} A_2 \cdot X - \frac{a_{21}}{a_{11}}A_1 \cdot X &= 0 \\ &\vdots \\ A_m \cdot X - \frac{a_{m1}}{a_{11}}A_1 \cdot X &= 0. \end{aligned}$$

En este sistema, los coeficientes de x_1 son iguales a 0. Por tanto, podemos considerar $(***)$ como un sistema de $m-1$ ecuaciones con $n-1$ incógnitas, y tenemos que $n - 1 > m - 1$.

Conforme a nuestra suposición, podemos hallar una solución no trivial $(x_1, \ldots, x_n)$ para este sistema. Entonces podemos despejar x_1 en la primera ecuación, o sea,

$$x_1 = \frac{-1}{a_{11}}(a_{12}x_2 + \cdots + a_{1n}x_n).$$

De esta manera, hallamos una solución de $A_1 \cdot X = 0$ pero, de acuerdo con $(***)$, tenemos

$$A_i \cdot X = \frac{a_{i1}}{a_{11}} A_1 \cdot X$$

para $i = 2, \ldots, m$. Por tanto, $A_i \cdot X = 0$ para $i = 2, \ldots, m$ y, por consiguiente, hemos encontrado una solución no trivial de nuestro sistema original $(**)$.

El argumento que acabamos de dar nos permite proceder por pasos de una ecuación a dos ecuaciones, luego de dos a tres, y así sucesivamente. Esto concluye la prueba.

Ejercicios II, §3

1. Sean

$$E_1 = (1, 0, \ldots, 0), \quad E_2 = (0, 1, 0, \ldots, 0), \quad \ldots, \quad E_n = (0, \ldots, 0, 1)$$

los vectores unitarios canónicos de $\mathbf{R}^n$. Sea X una n-tupla. Si $X \cdot E_i = 0$ para todo i, demuestre que $X = O$.

2. Sean $A_1, \ldots, A_m$ vectores de $\mathbf{R}^n$. Sean X y Y soluciones del sistema de ecuaciones

$$X \cdot A_i = 0 \quad \text{y} \quad Y \cdot A_i = 0 \quad \text{para} \quad i = 1, \ldots, m.$$

Demuestre que $X + Y$ también es una solución. Si c es un número, demuestre que cX es una solución.

3. En el ejercicio 2, suponga que X es perpendicular a cada uno de los vectores $A_1, \ldots, A_m$. Sean $c_1, \ldots, c_m$ números. Se dice que un vector

$$c_1 A_1 + \cdots + c_m A_m$$

es una combinación lineal de $A_1, \ldots, A_m$. Demuestre que X es perpendicular a dicho vector.

4. Considere el sistema no homogéneo $(*)$ que consiste en todos los X tales que $X \cdot A_i = b_i$ para $i = 1, \ldots, m$. Si X y X' son dos soluciones de este sistema, demuestre que existe una solución Y del sistema homogéneo $(**)$ tal que $X' = X + Y$. Recíprocamente, si X es cualquier solución de $(*)$ y Y es una solución de $(**)$, demuestre que $X + Y$ es una solución de $(*)$.

5. Halle al menos una solución no trivial para cada uno de los siguientes sistemas de ecuaciones. Como hay muchas posibilidades, no damos respuestas.

(a) $3x + y + z = 0$

(b) $$\begin{aligned} 3x + y + z &= 0 \\ x + y + z &= 0 \end{aligned}$$

(c) $$\begin{aligned} 2x - 3y + 4z &= 0 \\ 3x + y + z &= 0 \end{aligned}$$

(d) $$\begin{aligned} 2x + y + 4z + w &= 0 \\ -3x + 2y - 3z + w &= 0 \\ x + y + z &= 0 \end{aligned}$$

(e) $$\begin{aligned} -x + 2y - 4z + w &= 0 \\ x + 3y + z - w &= 0 \end{aligned}$$

(f) $$\begin{aligned} -2x + 3y + z + 4w &= 0 \\ x + y + 2z + 3w &= 0 \\ 2x + y + z - 2w &= 0 \end{aligned}$$

6. Demuestre que las únicas soluciones de los siguientes sistemas de ecuaciones son las triviales.

(a) $\begin{aligned} 2x + 3y &= 0 \\ x - y &= 0 \end{aligned}$

(b) $\begin{aligned} 4x + 5y &= 0 \\ -6x + 7y &= 0 \end{aligned}$

(c) $\begin{aligned} 3x + 4y - 2z &= 0 \\ x + y + z &= 0 \\ -x - 3y + 5z &= 0 \end{aligned}$

(d) $\begin{aligned} 4x - 7y + 3z &= 0 \\ x + y &= 0 \\ y - 6z &= 0 \end{aligned}$

(e) $\begin{aligned} 7x - 2y + 5z + w &= 0 \\ x - y + z &= 0 \\ y - 2z + w &= 0 \\ x + z + w &= 0 \end{aligned}$

(f) $\begin{aligned} -3x + y + z &= 0 \\ x - y + z - 2w &= 0 \\ x - z + w &= 0 \\ -x + y - 3w &= 0 \end{aligned}$

II, §4. Operaciones por renglones y eliminación de Gauss

Consideremos el sistema de ecuaciones lineales

$$\begin{aligned} 3x - 2y + z + 2w &= 1, \\ x + y - z - w &= -2, \\ 2x - y + 3z &= 4. \end{aligned}$$

La matriz de coeficientes es

$$\begin{pmatrix} 3 & -2 & 1 & 2 \\ 1 & 1 & -1 & -1 \\ 2 & -1 & 3 & 0 \end{pmatrix}.$$

Llamaremos **matriz aumentada** a la que obtenemos al insertar la columna

$$\begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}$$

como última columna, por lo que la matriz aumentada es

$$\begin{pmatrix} 3 & -2 & 1 & 2 & 1 \\ 1 & 1 & -1 & -1 & -2 \\ 2 & -1 & 3 & 0 & 4 \end{pmatrix}.$$

En general, sea $AX = B$ un sistema de m ecuaciones lineales con n incógnitas, que se escribe en detalle como sigue:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2, \\ \vdots \qquad\quad & \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_n, \end{aligned}$$

Entonces definimos la **matriz aumentada** como la matriz de m por $n + 1$:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{pmatrix}.$$

En los ejemplos de ecuaciones lineales homogéneas de la sección anterior, habrá observado que efectuamos las siguientes operaciones, conocidas como **operaciones elementales por renglones**:

Multiplicar una ecuación por un número no nulo.
Sumar una ecuación a otra.
Intercambiar dos ecuaciones.

Estas operaciones se reflejan en operaciones sobre la matriz de coeficientes aumentada, que también se conocen como **operaciones elementales por renglones:**

Multiplicar un renglón por un número no nulo.
Sumar un renglón a otro.
Intercambiar dos renglones.

Supongamos que se cambia un sistema de ecuaciones lineales mediante una operación elemental por renglones. Entonces las soluciones del nuevo sistema son exactamente las mismas que las soluciones del sistema original. Al hacer operaciones por renglones se espera simplificar la forma del sistema, de manera que sea más fácil encontrar las soluciones.

Decimos que dos matrices son **equivalentes por renglones** si una de ellas se puede obtener de la otra mediante una sucesión de operaciones elementales por renglones. Si A es la matriz de coeficientes de un sistema de ecuaciones lineales y B es el vector columna, como antes, de manera que

$$(A, B)$$

es la matriz aumentada, y si (A', B') es equivalente por renglones a (A, B), entonces las soluciones del sistema

$$AX = B$$

son las mismas que las soluciones del sistema

$$A'X = B'.$$

Con el objeto de obtener un sistema equivalente (A', B') tan sencillo como sea posible, usamos un método que primero ilustraremos en un caso concreto.

Ejemplo. Considere la matriz aumentada del ejemplo anterior. Tenemos las siguientes equivalencias por renglones:

$$\begin{pmatrix} 3 & -2 & 1 & 2 & 1 \\ 1 & 1 & -1 & -1 & -2 \\ 2 & -1 & 3 & 0 & 4 \end{pmatrix}$$

Reste 3 veces el segundo renglón del primero.

$$\begin{pmatrix} 0 & -5 & 4 & 5 & 7 \\ 1 & 1 & -1 & -1 & -2 \\ 2 & -1 & 3 & 0 & 4 \end{pmatrix}$$

Reste 2 veces el segundo renglón del tercero.

$$\begin{pmatrix} 0 & -5 & 4 & 5 & 7 \\ 1 & 1 & -1 & -1 & -2 \\ 0 & -3 & 5 & 2 & 8 \end{pmatrix}$$

Intercambie el primer renglón con el segundo; multiplique el segundo por -1.

$$\begin{pmatrix} 1 & 1 & -1 & -1 & -2 \\ 0 & 5 & -4 & -5 & -7 \\ 0 & -3 & 5 & 2 & 8 \end{pmatrix}$$

Multiplique el sengundo renglón por 3; multiplique el tercer renglón por 5.

$$\begin{pmatrix} 1 & 1 & -1 & -1 & -2 \\ 0 & 15 & -12 & -15 & -21 \\ 0 & -15 & 25 & 10 & 40 \end{pmatrix}$$

Sume el segundo renglón al tercero.

$$\begin{pmatrix} 1 & 1 & -1 & -1 & -2 \\ 0 & 15 & -12 & -15 & -21 \\ 0 & 0 & 13 & -5 & 19 \end{pmatrix}$$

Lo que hemos logrado es que cada renglón sucesivo tenga su primera componente no nula en al menos un lugar posterior con respecto al renglón anterior. Esto hace que sea muy sencillo resolver las ecuaciones. El nuevo sistema, cuya matriz aumentada es la última que obtuvimos, se puede escribir en la forma:

$$\begin{aligned} x + y - \quad z - \quad w &= -2, \\ 15y - 12z - 15w &= -21, \\ 13z - \quad 5w &= 19. \end{aligned}$$

Éste se encuentra ahora en una forma tal que podemos resolverlo dando a w un valor arbitrario en la tercera ecuación y despejar z a partir de la tercera ecuación. Luego despejamos y a partir de la segunda ecuación y x a partir de la primera. Con las fórmulas, esto da:

$$z = \frac{19 + 5w}{13},$$

$$y = \frac{-21 + 12z + 15w}{15},$$

$$x = -1 - y + z + w.$$

Para comenzar, podemos dar a w cualquier valor y después determinar los valores de x, y y z. Por tanto, vemos que las soluciones dependen de un parámetro libre. Más adelante expresaremos esta propiedad diciendo que el conjunto de soluciones tiene dimensión 1.

Por el momento, damos un nombre general al procedimiento anterior. Sea M una matriz. Decimos que M está en **forma escalonada por renglones** si tiene la siguiente propiedad:

Siempre que dos renglones sucesivos no contengan sólo ceros, entonces el segundo renglón comienza con una componente no nula al menos en un lugar posterior con respecto al primer renglón. Todos los renglones que sólo contengan ceros se encuentran en la parte inferior de la matriz.

En el ejemplo anterior transformamos una matriz en otra que se encuentra en forma escalonada por renglones. Los primeros coeficientes no nulos que aparecen a la izquierda en cada renglón se conocen como **coeficientes principales.** En el ejemplo anterior, los **coeficientes principales** son 1, 15, 13. Se puede efectuar un cambio adicional dividiendo cada renglón entre el coeficiente principal. Entonces la matriz anterior es equivalente por renglones a

$$\begin{pmatrix} 1 & 1 & -1 & -1 & -2 \\ 0 & 1 & \frac{2}{5} & \frac{1}{5} & 1 \\ 0 & 0 & 1 & -\frac{5}{13} & \frac{19}{13} \end{pmatrix}.$$

En esta última matriz, el coeficiente principal de cada renglón es igual a 1. Se podría hacer más operaciones por renglón para hacer aparecer más ceros; por ejemplo, restar el segundo renglón del primero y luego restar $\frac{2}{5}$ por el tercer renglón del segundo. Esto produce:

$$\begin{pmatrix} 1 & 0 & -\frac{7}{5} & -\frac{6}{5} & -3 \\ 0 & 1 & 0 & \frac{1}{5}+\frac{2}{13} & 1-\frac{38}{65} \\ 0 & 0 & 1 & -\frac{5}{13} & \frac{19}{13} \end{pmatrix}.$$

A menos que en la matriz resultante las fracciones no se vean tan horribles, por lo regular resulta engorroso hacer esta equivalencia adicional por renglón manualmente, aunque a una máquina no le importaría.

Ejemplo. La siguiente matriz se encuentra en una forma escalonada por renglones

$$\begin{pmatrix} 0 & 2 & -3 & 4 & 1 & 7 \\ 0 & 0 & 0 & 5 & 2 & -4 \\ 0 & 0 & 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Supongamos que esta matriz es la matriz aumentada de un sistema de ecuaciones lineales; entonces podemos resolver las ecuaciones lineales dando un valor arbitrario a algunas variables, tal como hicimos. En realidad, las ecuaciones son:

$$\begin{aligned} 2y - 3z + 4w + t &= 7, \\ 5w + 2t &= -4, \\ -3t &= 1. \end{aligned}$$

Entonces las soluciones son

$$\begin{aligned} t &= -1/3, \\ w &= \frac{-4-2t}{5}, \\ z &= \text{cualquier valor arbitrario dado}, \\ y &= \frac{7+3z-4w-t}{2}, \\ x &= \text{cualquier valor arbitrario dado}. \end{aligned}$$

El método de cambiar una matriz mediante equivalencias por renglones para llevarla a una forma escalonada funciona en general.

Teorema 4.1. *Toda matriz es equivalente por renglones a una matriz en forma escalonada por renglones.*

Demostración. Seleccionemos la primera componente no nula que aparezca a la izquierda en la matriz. Si esta componente no está en la primera columna, significa que la matriz sólo contiene ceros a la izquierda de esta componente y nos podemos olvidar de ellos. Por tanto, supongamos que esta componente no nula se encuentra en la primera columna. Después de un intercambio de renglones podemos encontrar una matriz equivalente tal que la esquina superior izquierda no sea 0. Digamos que la matriz es

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

y que $a_{11} \neq 0$. Multipliquemos el primer renglón por a_{21}/a_{11} y restémoslo del segundo. En forma análoga, multipliquemos el primer renglón por a_{i1}/a_{11} y restémoslo del i-ésimo renglón. Entonces obtenemos una matriz que tiene ceros en la primera columna excepto en a_{11}. Por tanto, la matriz original es equivalente por renglones a una matriz de la forma

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a'_{m2} & \cdots & a'_{mn} \end{pmatrix}.$$

Luego repetimos el procedimiento con la siguiente matriz de menor tamaño

$$\begin{pmatrix} a'_{22} & \cdots & a'_{2n} \\ \vdots & & \vdots \\ a'_{m2} & \cdots & a'_{mn} \end{pmatrix}.$$

Podemos continuar así hasta que la matriz esté en forma escalonada por renglones (formalmente mediante inducción). Esto concluye la demostración.

Observe que la demostración es sólo otra manera de formular el argumento de eliminación de la sección §3.

A continuación damos otra demostración del teorema fundamental:

Teorema 4.2. *Sea*

$$\begin{aligned} a_{11}x_1 + \cdots + \ & a_{1n}x_n = 0, \\ \vdots \qquad\quad & \vdots \\ a_{m1}x_1 + \cdots + \ & a_{mn}x_n = 0, \end{aligned}$$

un sistema de m ecuaciones lineales homogéneas con n incógnitas, donde $n > m$. Entonces existe una solución no trivial.

Demostración. Sea $A = (a_{ij})$ la matriz de coeficientes. Entonces A es equivalente a A', que se encuentra en forma escalonada por renglones:

$$\begin{aligned} a_{k_1}x_{k_1} + S_{k_1}(x) \qquad\qquad\quad &= 0, \\ a_{k_2}x_{k_2} + S_{k_2}(x) \qquad &= 0, \\ \cdots\cdots\cdots\cdots \\ a_{k_r}x_{k_r} + S_{k_r}(x) &= 0, \end{aligned}$$

donde $a_{k_1} \neq 0, \ldots, a_{k_r} \neq 0$ son los coeficientes no nulos de las variables que aparecen al principio a la izquierda en cada renglón sucesivo, y $S_{k_1}(x), \ldots, S_{k_r}(x)$ indican sumas de variables con ciertos coeficientes, pero tales que, si una variable x_j aparece en $S_{k_1}(x)$, entonces $j > k_1$, y lo mismo sucede con las otras sumas. Si x_j aparece en S_{k_i}, entonces $j > k_i$. Como, por hipótesis, el número total de variables n es estrictamente mayor que el número de ecuaciones, debemos tener $r < n$. En consecuencia, hay $n - r$ variables distintas de $x_{k_1}, \ldots, x_{k_r}$ y $n - r > 0$. Damos a estas variables valores arbitrarios y, por supuesto, los podemos seleccionar de manera que no todos sean iguales a 0. Luego despejamos las variables $x_{k_r}, x_{k_{r-1}}, \ldots, x_{k_1}$, comenzando con la ecuación inferior y prosiguiendo hacia arriba, por ejemplo,

$$\begin{aligned} x_{k_r} &= -S_{k_r}(x)/a_{k_r}, \\ x_{k_{r-1}} &= -S_{k_{r-1}}(x)/a_{k_{r-1}}, \quad \text{y así sucesivamente.} \end{aligned}$$

Esto nos da la solución no trivial y así el teorema queda probado.

Observe que el patrón se ajusta exactamente al de los ejemplos, aunque con una notación que se refiere al caso general.

Ejercicios II, §4

En cada uno los siguientes casos, encuentre una matriz equivalente por renglones que se encuentre en la forma escalonada por renglones.

1. (a) $\begin{pmatrix} 6 & 3 & -4 \\ -4 & 1 & -6 \\ 1 & 2 & -5 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix}$.

2. (a) $\begin{pmatrix} 1 & -2 & 3 & -1 \\ 2 & -1 & 2 & 2 \\ 3 & 1 & 2 & 3 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 1 & 3 & -2 \\ 2 & 1 & -4 & 3 \\ 2 & 3 & 2 & -1 \end{pmatrix}$.

3. (a) $\begin{pmatrix} 1 & 2 & -1 & 2 & 1 \\ 2 & 4 & 1 & -2 & 3 \\ 3 & 6 & 2 & -6 & 5 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 3 & -1 & 2 \\ 0 & 11 & -5 & 3 \\ 2 & -5 & 3 & 1 \\ 4 & 1 & 1 & 5 \end{pmatrix}$.

4. Escriba las matrices de coeficientes de las ecuaciones lineales del ejercicio 5 de la sección §3, y en cada caso muestre una matriz equivalente por renglones que se encuentre en forma escalonada. En cada caso, resuelva las ecuaciones por este método.

II, §5. Operaciones por renglones y matrices elementales

Antes de leer esta sección, resuelva los ejemplos numéricos que aparecen en los ejercicios 33 a 37 de la sección §2.

Las operaciones por renglones que usamos para resolver ecuaciones lineales se pueden representar mediante operaciones con matrices. Sean $1 \leq r \leq m$ y $1 \leq s \leq m$. Sea I_{rs} la matriz cuadrada de $m \times m$ que tiene por componente a 1 en el lugar rs y 0 en los demás:

$$I_{rs} = \begin{pmatrix} 0 \cdots\cdots\cdots 0 \\ \vdots \qquad\quad \vdots \\ 0 \cdots 1_{rs} \cdots 0 \\ \vdots \qquad\quad \vdots \\ 0 \cdots\cdots\cdots 0 \end{pmatrix}.$$

Sea $A = (a_{ij})$ cualquier matriz de $m \times n$. ¿Cuál es el efecto de la multiplicación $I_{rs}A$?

$$r\left\{\begin{pmatrix} 0 \cdots\cdots\cdots 0 \\ \vdots \qquad\quad \vdots \\ 0 \cdots 1_{rs} \cdots 0 \\ \vdots \qquad\quad \vdots \\ 0 \cdots\cdots\cdots 0 \end{pmatrix}\right. \!\!\!\!\!\!\!\!\!\!\!\!\underbrace{}_{s} \begin{pmatrix} a_{11} \cdots a_{1n} \\ \vdots \qquad \vdots \\ \vdots \qquad \vdots \\ a_{s1} \cdots a_{sn} \\ \vdots \qquad \vdots \\ a_{n1} \cdots a_{nn} \end{pmatrix} \Big\} s = \begin{pmatrix} 0 \quad \cdots 0 \\ \vdots \qquad \vdots \\ a_{s1} \cdots a_{sn} \\ \vdots \qquad \vdots \\ 0 \quad \cdots 0 \end{pmatrix} \Big\} r.$$

La definición de multiplicación de matrices muestra que $I_{rs}A$ es la matriz obtenida al poner el renglón s de A en el renglón r y ceros en los demás lugares.

Si $r = s$, entonces I_{rr} tiene una componente 1 sobre el lugar de la diagonal y cero en los demás sitios. La multiplicación por I_{rr} deja fijo entonces al renglón r y reemplaza todos los demás renglones por ceros.

Si $r \neq s$, sea

$$J_{rs} = I_{rs} + I_{sr}.$$

Entonces

$$J_{rs}A = I_{rs}A + I_{sr}A.$$

Luego, $I_{rs}A$ pone al renglón s de A en el lugar r e $I_{sr}A$ pone el renglón r de A en el lugar s. Todos los otros renglones se reemplazan por ceros. Así, J_{rs} intercambia al renglón r con el s y reemplaza todos los demás renglones por ceros.

Ejemplo. Sean

$$J = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{y} \quad A = \begin{pmatrix} 3 & 2 & -1 \\ 1 & 4 & 2 \\ -2 & 5 & 1 \end{pmatrix}.$$

Si el lector efectúa la multiplicación de matrices, verá directamente que JA intercambia los renglones primero y segundo de A y reemplaza al tercer renglón por ceros.

Por otro lado, sea

$$E = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Entonces EA es la matriz obtenida de A al intercambiar el primer renglón con el segundo y al dejar fijo el tercer renglón. Podemos expresar E como una suma:

$$E = I_{12} + I_{21} + I_{33}$$

donde I_{rs} es la matriz que tiene 1 en la componente rs y, como antes, 0 en las demás componentes. Observe que E se obtiene a partir de la matriz unitaria intercambiando los dos primeros renglones y dejando fijo el tercer renglón. Así, la operación de intercambiar los dos primeros renglones de A se lleva a cabo mediante la multiplicación con la matriz E obtenida al hacer esta operación sobre la matriz unitaria.

Éste es un caso especial del siguiente hecho general.

Teorema 5.1. *Sea E la matriz obtenida a partir de la matriz unitaria de $n \times n$ al intercambiar dos renglones. Sea A una matriz de $n \times n$. Entonces EA es la matriz obtenida a partir de A al intercambiar estos dos renglones.*

Demostración. La demostración se lleva a cabo de acuerdo al patrón del ejemplo; sólo es cuestión de determinar cuáles son los símbolos que se van a usar.

Suponga que intercambiamos el renglón r con el renglón s. Entonces podemos escribir

$$E = I_{rs} + I_{sr} + \text{suma de las matrices } I_{jj} \text{ con } j \neq r, j \neq s.$$

Así, E difiere de la matriz unitaria en que se intercambiaron los renglones r y s. Entonces

$$EA = I_{rs}A + I_{sr}A + \text{suma de las matrices } I_{jj}A,$$

donde $j \neq r$, $j \neq s$. Como consecuencia de nuestro análisis anterior, ésta es precisamente la matriz obtenida al intercambiar los renglones r y s de A y al dejar fijas todas las demás componentes.

El mismo tipo de análisis se aplica al siguiente resultado.

Teorema 5.2. *Sea E la matriz obtenida de la matriz unitaria de $n \times n$ al multiplicar el renglón r por un número c y sumarlo al renglón s, $r \neq s$. Sea A una matriz de $n \times n$. Entonces EA se obtiene a partir de A al multiplicar el renglón r de A por c y al sumarlo al renglón s de A.*

Demostración. Podemos escribir

$$E + I + cI_{sr}.$$

Entonces $EA = A + cI_{sr}A$. Sabemos que $I_{sr}A$ pone el renglón r de A en el lugar s y, al multiplicar por c, se multiplica este renglón por c. Todos los renglones, aparte del renglón s de $cI_{sr}A$, son iguales a 0. Por consiguiente, la suma $A + cI_{sr}A$ equivale a sumar c veces el renglón r de A al renglón s de A, como se quería mostrar.

Ejemplo. Sea

$$E = \begin{pmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Entonces E se obtiene de la matriz unitaria sumando 4 veces el tercer renglón al primero. Considere cualquier matriz A de $4 \times n$ y calcule EA. Encontrará que EA se obtiene al multiplicar el tercer renglón de A por 4 y sumarlo al primer renglón de A.

En forma más general, podemos suponer que $E_{rs}(c)$ para $r \neq s$ es la matriz elemental.

$$E_{rs}(c) = I + cI_{rs}.$$

$$r\left\{ \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 \cdots c \cdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{pmatrix} \right. \quad (\text{columna } s \text{ marcada arriba})$$

Difiere de la matriz unitaria en que tiene la componente rs igual a c. El efecto de multiplicar por la izquierda por $E_{rs}(c)$ consiste en sumar c veces el renglón s al renglón r.

Una **matriz elemental** será una matriz de cualquiera de los siguientes tres tipos:

(a) Una matriz que se obtiene de la matriz unitaria al multiplicar la r-ésima componente de la diagonal por un número $c \neq 0$.

(b) Una matriz que se obtiene de la matriz unitaria al intercambiar dos renglones (digamos el renglón r con el renglón s, $r \neq s$).

(c) Una matriz $E_{rs}(c) = I + cI_{rs}$, con $r \neq s$, y que tiene a c como su componente rs cuando $r \neq s$, y a todas las demás componentes iguales a 0, excepto las componentes diagonales que son iguales a 1.

En estos tres tipos se reflejan las operaciones por renglones que estudiamos en la sección anterior.

La multiplicación por una matriz del tipo (a) multiplica el renglón r por el número c.

La multiplicación por una matriz del tipo (b) intercambia el renglón r con el s.

La multiplicación por una matriz del tipo (c) suma c veces el renglón s al r.

Proposición 5.3. *Una matriz elemental es invertible.*

Demostración. Con respecto al tipo (a), la matriz inversa tiene por r-ésima componente diagonal a c^{-1}, debido a que, al multiplicar un renglón primero por c y luego por c^{-1}, el renglón no cambia.

Con respecto al tipo (b), observamos que, al intercambiar el renglón r con el s dos veces, se vuelve a la misma matriz con la que comenzamos.

Con respecto al tipo (c), igual que en el teorema 5.2, sea E la matriz que suma c veces el renglón s al r de la matriz unitaria. Sea D la matriz que suma $-c$ veces el renglón s al r de la matriz unitaria (para $r \neq s$). Entonces DE es la matriz unitaria, y por lo tanto, también lo es ED, de manera que E es invertible.

Ejemplo. Las siguientes matrices elementales son inversas entre sí:

$$E = \begin{pmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad E^{-1} = \begin{pmatrix} 1 & 0 & -4 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Mostraremos una manera efectiva de encontrar la inversa de una matriz cuadrada, si es que tiene. Esto se basará en las siguientes propiedades.

Si A y B son matrices cuadradas del mismo tamaño y tienen inversas, entonces también tiene inversa AB y

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Esto es inmediato, pues

$$ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I.$$

En forma análoga, para cualquier cantidad de factores, se tiene:

Proposición 5.4. *Si $A_1, \ldots, A_k$ son matrices invertibles del mismo tamaño, entonces su producto tiene una inversa y*

$$(A_1, \ldots, A_k)^{-1} = A_k^{-1} \cdots A_1^{-1}.$$

Observe que, en el miembro derecho de la igualdad, tomamos el producto de las inversas en orden inverso. Entonces

$$A_1 \cdots A_k A_k^{-1} \cdots A_1 = I$$

debido a que podemos reducir $A_k A_k^{-1}$ a I, luego $A_{k-1}A_{k-1}^{-1}$ a I, y así sucesivamente.

Como una matriz elemental tiene inversa, concluimos que cualquier producto de matrices elementales tiene una inversa.

Proposición 5.5. *Sea A una matriz cuadrada y sea A' equivalente por renglones a A. Entonces A tiene una inversa si, y sólo si, A' tiene una inversa.*

Demostración. Existen matrices elementales $E_1, \ldots, E_k$ tales que

$$A' = E_1 \cdots E_k A.$$

Suponga que A tiene inversa. Entonces el miembro derecho de la igualdad tiene una inversa debido a la Proposición 5.4, puesto que el miembro derecho de la igualdad es un producto de matrices invertibles. En consecuencia, A' tiene una inversa. Esto prueba la proposición.

Ahora estamos en posibilidades de encontrar una inversa para una matriz cuadrada A, si es que tiene. Por el Teorema 4.1 sabemos que A es equivalente por renglones a una matriz A' en forma escalonada. Si un renglón de A' es igual a cero, entonces, por la definición de forma escalonada, el último renglón debe ser igual a cero y A' no es invertible; en consecuencia, A no es invertible. Si todos los renglones de A' son no nulos, entonces A' es una matriz triangular con sus componentes diagonales no nulas. Ahora es suficiente encontrar una inversa para dicha matriz. En efecto, probamos:

Teorema 5.6. *Una matriz cuadrada A es invertible si, y sólo si, A es equivalente por renglones a la matriz unitaria. Cualquier matriz superiormente triangular cuyos elementos no son nulos es invertible.*

Demostración. Supongamos que A es equivalente por renglones a la matriz unitaria. Entonces A es invertible debido a la Proposición 5.5. Suponga que A es invertible; acabamos de ver que A es equivalente por renglones a una matriz

superiormente triangular cuyos elementos diagonales son no nulos. Suponga que A es dicha matriz:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Por suposición tenemos que $a_{11} \cdots a_{nn} \neq 0$. Multipliquemos el renglón i por a_{ii}^{-1}. Obtenemos una matriz triangular tal que todas las componentes diagonales son iguales a 1. Así, para probar el teorema, es suficiente hacerlo para este caso y podemos suponer que A tiene la forma

$$\begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 0 & 1 & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Multipliquemos el último renglón por a_{in} y restémoslo del renglón i para $i = 1, \ldots, n-1$. Esto hace que todos los elementos de la última columna sean iguales a 0 excepto el que se encuentra en la esquina inferior de la izquierda y que es 1. Repitamos este procedimiento con el que se encuentra próximo al último renglón y continuemos hacia arriba. Esto significa que, mediante equivalencias por renglones, podemos reemplazar por cero todas las componentes que se encuentran estrictamente arriba de la diagonal. Luego terminamos con la matriz unitaria que, por consiguiente, es equivalente por renglones a la matriz original. Esto prueba el teorema.

Corolario 5.7. *Sea A una matriz invertible. Entonces A se puede expresar como producto de matrices elementales.*

Demostración. Esto se debe a que A es equivalente por renglones a la matriz unitaria y a que las operaciones por renglones se representan mediante la multiplicación por matrices elementales, de manera que existen $E_1, \ldots, E_k$ tales que

$$E_k \cdots E_1 A = I.$$

Entonces $A = E_1^{-1} \cdots E_k^{-1}$, con lo que se prueba el corolario.

Cuando A queda expresada de esa manera, también obtenemos una expresión para la inversa de A, a saber,

$$A^{-1} = E_k \cdots E_1.$$

Las matrices elementales $E_1 \ldots E_k$ son aquellas que se usan para convertir A en la matriz unitaria.

Ejemplo. Sean

$$A = \begin{pmatrix} 2 & -3 & 1 \\ 1 & 1 & -1 \\ 2 & 0 & 1 \end{pmatrix}, \qquad I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Queremos encontrar una inversa para A. Efectuamos las siguientes operaciones por renglones, las cuales corresponden a la multiplicación por las matrices elementales que se muestran.

Intercambie los primeros dos renglones.

$$\begin{pmatrix} 1 & 1 & -1 \\ 2 & -3 & 1 \\ 2 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Reste 2 veces el primer renglón del segundo.
Reste 2 veces el primer renglón del tercero.

$$\begin{pmatrix} 1 & 1 & -1 \\ 0 & -5 & 3 \\ 0 & -2 & 3 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & -2 & 1 \end{pmatrix}.$$

Reste el segundo renglón, multiplicado por 2/5, del tercero.

$$\begin{pmatrix} 1 & 1 & -1 \\ 0 & -5 & 3 \\ 0 & 0 & 9/5 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 & 0 \\ 1 & -2 & 0 \\ -2/5 & -6/5 & 1 \end{pmatrix}.$$

Reste el tercer renglón, multiplicado por 5/3, del segundo.
Sume al primer renglón el tercero multiplicado por 5/9.

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 9/5 \end{pmatrix}, \qquad \begin{pmatrix} -2/9 & 1/3 & 5/9 \\ 5/3 & 0 & -5/3 \\ -2/5 & -6/5 & 1 \end{pmatrix}.$$

Sume al primer renglón el segundo multiplicado por 1/5.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 9/5 \end{pmatrix}, \qquad \begin{pmatrix} 1/9 & 1/3 & 2/9 \\ 5/3 & 0 & -5/3 \\ -2/5 & -6/5 & 1 \end{pmatrix}.$$

Multiplique el segundo renglón por $-1/5$.
Multiplique el tercer renglón por 5/9.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1/9 & 1/3 & 2/9 \\ -1/3 & 0 & 1/3 \\ -2/9 & -2/3 & 5/9 \end{pmatrix}.$$

Entonces A^{-1} es la matriz de la derecha, esto es,

$$A^{-1} = \begin{pmatrix} 1/9 & 1/3 & 2/9 \\ -1/3 & 0 & 1/3 \\ -2/9 & -2/3 & 5/9 \end{pmatrix}.$$

El lector puede verificar esta afirmación multiplicando directamente por A con el objeto de hallar la matriz unitaria.

Si A es una matriz cuadrada y consideramos un sistema no homogéneo de ecuaciones lineales

$$AX = B,$$

entonces, si A es invertible, podemos usar la inversa para resolver el sistema. En efecto, en este caso, multiplicamos por la izquierda ambos miembros de la igualdad por A^{-1} y hallamos que

$$X = A^{-1}B.$$

Esto también prueba lo siguiente:

Proposición 5.8. *Sea $AX = B$ un sistema de n ecuaciones lineales con n incógnitas. Suponga que la matriz de coeficientes A es invertible. Entonces existe una solución única X del sistema y*

$$X = A^{-1}B.$$

Ejercicios II, §5

1. Mediante el uso de operaciones elementales por renglones, halle inversas para las siguientes matrices.

(a) $\begin{pmatrix} 2 & 1 & 2 \\ 0 & 3 & -1 \\ 4 & 1 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 3 & -1 & 5 \\ -1 & 2 & 1 \\ -2 & 4 & 3 \end{pmatrix}$

(c) $\begin{pmatrix} 2 & 4 & 3 \\ -1 & 3 & 0 \\ 0 & 2 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 0 & 2 & 7 \end{pmatrix}$

(e) $\begin{pmatrix} -1 & 5 & 3 \\ 4 & 0 & 0 \\ 2 & 7 & 8 \end{pmatrix}$ (f) $\begin{pmatrix} 3 & 1 & 2 \\ 4 & 5 & 1 \\ -1 & 2 & -3 \end{pmatrix}$

Nota: En el capítulo sobre determinantes se encontrará otra manera de hallar inversas.

2. Sea $r \neq s$. Demuestre que $I_{rs}^2 = O$.

3. Sea $r \neq s$. Sea $E_{rs}(c)$ tal como aparece en el texto. Demuestre que

$$E_{rs}(c)E_{rs}(c') = E_{rs}(c + c').$$

II, §6. Combinaciones lineales

Sean $A^1, \ldots, A^n$ m-tuplas de $\mathbf{R}^m$. Sean $x_1, \ldots, x_n$ números. Entonces decimos que

$$x_1A^1 + \cdots + x_nA^n$$

es una **combinación lineal** de $A^1, \ldots, A^n$; $x_1, \ldots, x_n$ se denominan **coeficientes** de la combinación lineal. Se aplica una definición similar a una combinación lineal de vectores renglón.

Se dice que la combinación lineal es **no trivial** si no todos los coeficientes $x_1, \ldots, x_n$ son iguales a 0.

Una vez más, consideremos un sistema de ecuaciones lineales homogéneas

$$(**) \qquad \begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ \vdots \qquad\qquad & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0. \end{aligned}$$

Nuestro sistema de ecuaciones homogéneas también se puede escribir en la forma siguiente:

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

o, en forma más concisa:

$$x_1A^1 + \cdots + x_nA^n = O,$$

donde $A^1, \ldots, A^n$ son los vectores columna de la matriz de coeficientes $A = (a_{ij})$. Así, el problema de hallar una solución no trivial del sistema de ecuaciones lineales homogéneas es equivalente a hallar una combinación lineal no trivial de $A^1, \ldots, A^n$ que sea igual a O.

Se dice que los vectores $A^1, \ldots, A^n$ son **linealmente dependientes** si existen números $x_1, \ldots, x_n$, no todos iguales a 0, tales que

$$x_1A^1 + \cdots + x_nA^n = O.$$

De este modo, una solución no trivial $(x_1, \ldots, x_n)$ es una n-tupla que da una combinación lineal de $A^1, \ldots, A^n$ igual a O, es decir, una relación de dependencia lineal entre las columnas de A. *Así pues, podemos resumir la descripción del conjunto de soluciones del sistema de ecuaciones lineales homogéneas en una tabla.*

(a) *Consiste en aquellos vectores* X *que dan relaciones lineales*

$$x_1A^1 + \cdots + x_nA^n = O$$

entre las columnas de A.

(b) *Consiste en aquellos vectores* X *que son perpendiculares a los renglones de* A, *esto es,* $X \cdot A_i = 0$ *para todo* i.

(c) *Consiste en aquellos vectores* X *tales que* $AX = O$.

Se dice que los vectores $A^1, \ldots, A^n$ son **linealmente independientes** si, dada cualquier combinación lineal de ellos que sea igual a O, esto es,

$$x_1A^1 + \cdots + x_nA^n = O,$$

entonces necesariamente debemos tener que $x_j = 0$ para todo $j = 1, \ldots, n$. Esto significa que no hay relaciones no triviales de dependencia lineal entre los vectores $A^1, \ldots, A^n$.

Ejemplo. Los vectores unitarios canónicos

$$E_1 = (1, 0, \ldots, 0), \ldots, E_n = (0, \ldots, 0, 1)$$

de $\mathbf{R}^n$ son linealmente independientes. En efecto, sean $x_1, \ldots, x_n$ números tales que

$$x_1 E^1 + \cdots + x_n E^n = O.$$

El miembro de la izquierda de la igualdad es, precisamente, la n-tupla $(x_1, \ldots, x_n)$. Si esta n-tupla es O, entonces todas las componentes son 0, de manera que $x_i = 0$ para todo i. Esto prueba que $E_1, \ldots, E_n$ son linealmente independientes.

En el siguiente capítulo estudiaremos las nociones de dependencia e independencia lineales en forma más sistemática. Se mencionaron en esta parte sólo para tener una tabla completa para las tres interpretaciones básicas de un sistema de ecuaciones lineales, y para introducir la noción en un caso concreto especial antes de dar las definiciones generales en espacios vectoriales.

Ejercicio II, §6

1. (a) Sean $A = (a_{ij})$ y $B = (b_{jk})$, y sea $AB = C$, donde $C = (c_{ik})$. Sea C^k la k-ésima columna de C. Exprese C^k como una combinación lineal de las columnas de A. Describa con precisión cuáles son los coeficientes que provienen de la matriz B.
   (b) Sea $AX = C^k$, donde X es alguna columna de B. ¿De qué columna se trata?

CAPÍTULO III

Espacios vectoriales

Igual que siempre, una colección de objetos se llamará **conjunto**. Un miembro del conjunto también se conoce como **elemento** del conjunto. En la práctica resulta útil emplear símbolos breves para denotar ciertos conjuntos. Por ejemplo, denotamos con $\mathbf{R}$ el conjunto de todos los números. Decir que "x es un número" o que "x es un elemento de $\mathbf{R}$" es lo mismo. Con $\mathbf{R}^n$ se denotará el conjunto de n-tuplas de números. Así, "x es un elemento de $\mathbf{R}^n$" y "x es una n-tupla" significan lo mismo. En vez de decir que u es un elemento de un conjunto S, con frecuencia decimos que u **pertenece a** S, y escribimos $u \in S$. Si S y S' son dos conjuntos y si todo elemento de S' es un elemento de S, entonces decimos que S' es un **subconjunto** de S. Así, el conjunto de números racionales es un subconjunto del conjunto de números (reales). Decir que S es un subconjunto de S' es decir que S es parte de S'. Para denotar el hecho de que S es un subconjunto de S', escribimos $S \subset S'$.

Si S_1 y S_2 son conjuntos, entonces la **intersección** de S_1 y S_2, que se denota con $S_1 \cap S_2$, es el conjunto de elementos que pertenecen tanto a S_1 como a S_2. La **unión** de S_1 y S_2, que se denota con $S_1 \cup S_2$, es el conjunto de elementos que pertenecen a S_1 o a S_2.

III, §1. Definiciones

En matemáticas nos encontramos con varios tipos de objetos que se pueden sumar y multiplicar por números. Entre éstos se encuentran los vectores (de la misma dimensión) y las funciones. Ahora resulta conveniente definir en general una noción que los incluya como un caso especial.

Un **espacio vectorial** V es un conjunto de objetos que se pueden sumar y multiplicar por números, de tal manera que la suma de dos elementos de V es de nuevo un elemento de V, el producto de un elemento de V por un número es un elemento de V, y se satisfacen las siguientes propiedades:

EV 1. *Dados los elementos u, v y w de V, tenemos*

$$(u+v)+w = u+(v+w).$$

EV 2. *Existe un elemento de V, que se denota con O, tal que*

$$O+u = u+O = u$$

para todos los elementos u de V.

EV 3. *Dado un elemento u de V, el elemento $(-1)u$ es tal que*

$$u+(-1)u = O.$$

EV 4. *Para todos los elementos u y v de V, tenemos*

$$u+v = v+u.$$

EV 5. *Si c es un número, entonces $c(u+v) = cu+cv$.*

EV 6. *Si a y b son dos números, entonces $(a+b)v = av+bv$.*

EV 7. *Si a y b son dos números, entonces $(ab)v = a(bv)$.*

EV 8. *Para todos los elementos u de V, tenemos que $1 \cdot u = u$ (en este caso 1 es el número uno).*

Hemos empleado todas estas reglas al trabajar con vectores o con funciones, pero deseamos ser más sistemáticos a partir de ahora, por lo que hemos hecho una lista de tales reglas. En los ejercicios aparecen más propiedades, las cuales se pueden deducir fácilmente de las que aparecen en esta lista y de ahora en adelante se darán por conocidas.

Las propiedades algebraicas de los elementos de un espacio vectorial arbitrario son muy semejantes a las de los elementos de $\mathbf{R}^2$, $\mathbf{R}^3$ o $\mathbf{R}^n$. En consecuencia, se acostumbra denominar **vectores** también a los elementos de un espacio vectorial arbitrario.

Si u y v son vectores (esto es, elementos del espacio vectorial arbitrario V), entonces la suma

$$u+(-1)v$$

usualmente se escribe $u-v$. También escribimos $-v$ en lugar de $(-1)v$.

Ejemplo 1. Fije dos enteros positivos m y n. Sea V el conjunto de todas las matrices de $m \times n$. También denotamos V con $\mathrm{Mat}(m \times n)$. Entonces V es un espacio vectorial. Es fácil verificar que mediante nuestras reglas para la adición de matrices y para la multiplicación de matrices por números se satisfacen todas las propiedades de la **EV 1** a la **EV 8**. El hecho principal que hay que observar en este caso es que la adición de matrices está definida en términos de las componentes, y que para la adición de componentes se satisfacen las condiciones análogas a las **EV 1** a **EV 4**. Son propiedades estándar de los números. Del mismo modo, las propiedades **EV 5** a **EV 8** son ciertas para la multiplicación de matrices por números, debido a que son ciertas las correspondientes propiedades para la multiplicación de números.

Ejemplo 2. Sea V el conjunto de todas las funciones definidas para todos los números. Si f y g son dos funciones, entonces sabemos cómo formar su suma $f + g$. Es la función cuyo valor en un número t es $f(t) + g(t)$. También sabemos cómo multiplicar f por un número c. Es la función cf cuyo valor en un número t es $cf(t)$. Al trabajar con funciones, hemos usado repetidas veces las propiedades **EV 1** a **EV 8**. Ahora nos hemos dado cuenta de que el conjunto de funciones es un espacio vectorial.

La función f tal que $f(t) = 0$ **para todo** t es la **función nula**. Insistimos en que la condición es para todo t. Si una función tiene algunos de sus valores iguales a cero, pero otros valores no son iguales a 0, entonces **no** es la función nula.

En la práctica, cierto número de propiedades elementales concernientes a la adición de elementos en un espacio vectorial resulta evidente, debido a la forma concreta en que está dado el espacio vectorial en términos de números, tal como se aprecia en los dos ejemplos anteriores. Ahora veremos en forma breve cómo probar tales propiedades a partir de los axiomas.

Es posible sumar varios elementos de un espacio vectorial. Suponga que queremos sumar cuatro elementos, digamos u, v, w y z. Primero sumamos dos cualesquiera de ellos, luego un tercero y finalmente un cuarto elemento. Al usar las reglas **EV 1** a **EV 4** vemos que no importa en qué orden efectuemos las sumas. Ésta es exactamente la misma situación que teníamos con los vectores. Por ejemplo, tenemos

$$\begin{aligned} ((u + v) + w) + z &= (u + (v + w)) + z \\ &= ((v + w) + u) + z \\ &= (v + w) + (u + z), \quad \text{etc.} \end{aligned}$$

Por ello se acostumbra eliminar los paréntesis y simplemente se escribe

$$u + v + w + z.$$

Se aplica la misma observación a la suma de cualquier número n de elementos de V.

Usamos 0 para denotar el número cero y O para denotar el elemento de cualquier espacio vectorial V que satisfaga la propiedad **EV 2**. También lo llamamos cero, pero no hay posibilidad alguna de confusión. Observemos que el elemento cero O queda determinado en forma única por la condición **EV 2**. En efecto, si

$$v + w = v,$$

entonces, al sumar $-v$ en ambos lados, se obtiene

$$-v + v + w = -v + v = O,$$

y el miembro de la izquierda de la igualdad es $O + w = w$, de manera que $w = O$.

Observe que para cualquier elemento v de V, tenemos

$$0v = O.$$

Demostración.

$$O = v + (-1)v = (1 - 1)v = 0v.$$

En forma análoga, si c es un número, entonces

$$cO = O.$$

Demostración. Tenemos que $cO = c(O + O) = cO + cO$. Si sumamos $-cO$ a ambos miembros de la igualdad obtenemos $cO = O$.

Subespacios

Sea V un espacio vectorial y sea W un subconjunto de V. Suponga que W satisface las siguientes condiciones.

(i) Si v y w son elementos de W, su suma $v + w$ también es un elemento de W.
(ii) Si v es un elemento de W y c es un número, entonces cv es un elemento de W.
(iii) El elemento O de V también es un elemento de W.

Entonces el propio W es un espacio vectorial. En efecto, las propiedades **EV 1** a **EV 8**, al ser satisfechas por todos los elementos de V, también lo son por los elementos de W. Decimos que W es un **subespacio** de V.

Ejemplo 3. Sea $V = \mathbf{R}^n$ y sea W el conjunto de vectores de V cuya última coordenada es igual a 0. Entonces W es un subespacio de V, al que podríamos identificar con $\mathbf{R}^{n-1}$.

Ejemplo 4. Sea A un vector de $\mathbf{R}^3$. Sea W el conjunto de todos los elementos B de $\mathbf{R}^3$ tales que $B \cdot A = 0$, esto es, tales que B es perpendicular a A. Entonces W es un subespacio de $\mathbf{R}^3$. Para ver esto observe que $O \cdot A = 0$, de manera que O está en W. Luego, suponga que B y C son perpendiculares a A. Entonces

$$(B + C) \cdot A = B \cdot A + C \cdot A = 0,$$

por lo que $B + C$ también es perpendicular a A. Por último, si x es un número, entonces

$$(xB) \cdot A = x(B \cdot A) = 0,$$

de manera que xB es perpendicular a A. Esto prueba que W es un subespacio de $\mathbf{R}^3$.

En forma más general, si A es un vector de $\mathbf{R}^n$, entonces el conjunto de todos los elementos B de $\mathbf{R}^n$ tales que $B \cdot A = 0$ es un subespacio de $\mathbf{R}^n$. La prueba es la misma que en el caso en que $n = 3$.

Ejemplo 5. Sea $\mathrm{Sim}(n \times n)$ el conjunto de todas las matrices simétricas de $n \times n$. Entonces $\mathrm{Sim}(n \times n)$ es un subespacio del espacio de todas las matrices de $n \times n$. En efecto, Si A y B son simétricas y c es un número, entonces $A + B$ y cA son simétricas. La matriz nula también es simétrica.

Ejemplo 6. Si f y g son dos funciones continuas, entonces $f + g$ es continua. Si c es un número, entonces cf es continua. La función nula es continua. En consecuencia, las funciones continuas forman un subespacio del espacio vectorial de todas las funciones.

Si f y g son dos funciones diferenciables, entonces su suma $f + g$ es diferenciable. Si c es un número, entonces cf es diferenciable. La función nula es diferenciable. Por tanto, las funciones diferenciables forman un subespacio del espacio vectorial de todas las funciones. Además, toda función diferenciable es continua. Por consiguiente, las funciones diferenciables forman un subespacio del espacio vectorial de las funciones continuas.

Ejemplo 7. Sea V un espacio vectorial y sean U y W subespacios. Denotamos con $U \cap W$ la intersección de U y W, esto es, el conjunto de elementos que pertenecen tanto a U como a W. Entonces $U \cap W$ es un subespacio. Por ejemplo, si U y W son dos planos en el espacio de 3 dimensiones que pasan por el origen, entonces su intersección, en general, será una recta que pasa por el origen, tal como se muestra en la figura 1.

Ejemplo 8. Sean U y W subespacios de un espacio vectorial V. Denotamos con

$$U + W$$

el conjunto de todos los elementos $u + w$, donde $u \in U$ y $w \in W$. Dejamos que el lector verifique que $U + W$ es un subespacio de V, el que se dice está **generado por** $U + W$ y que se conoce como la **suma** de U y W.

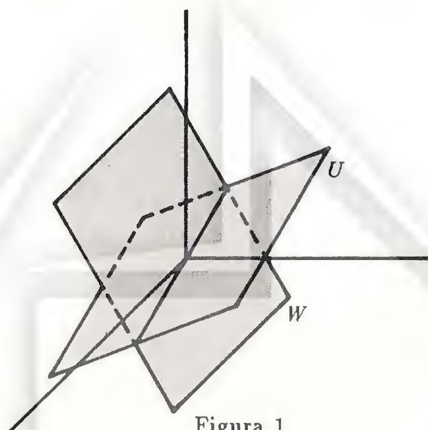


Figura 1

Ejercicios III, §1

1. Sean $A_1, \ldots, A_r$ vectores de $\mathbf{R}^n$. Sea W el conjunto de los vectores B de $\mathbf{R}^n$ tales que $B \cdot A_i = 0$ para todo $i = 1, \ldots, r$. Muestre que W es un subespacio de $\mathbf{R}^n$.

2. Muestre que los siguientes conjuntos de elementos de $\mathbf{R}^2$ forman subespacios.
   (a) El conjunto de todas las (x, y) tales que $x = y$.
   (b) El conjunto de todas las (x, y) tales que $x - y = 0$.
   (c) El conjunto de todas las (x, y) tales que $x + 4y = 0$.

3. Muestre que los siguientes conjuntos de elementos de $\mathbf{R}^3$ forman subespacios.
   (a) El conjunto de todas las (x, y, z) tales que $x + y + z = 0$.
   (b) El conjunto de todas las (x, y, z) tales que $x = y$ y $2y = z$.
   (c) El conjunto de todas las (x, y, z) tales que $x + y = 3z$.

4. Si U y W son subespacios de un espacio vectorial V, entonces muestre que $U \cap W$ y $U + W$ son subespacios.

5. Sea V un subespacio de $\mathbf{R}^n$. Sea W el conjunto de elementos de $\mathbf{R}^n$ que son perpendiculares a todo elemento de V. Demuestre que W es un subespacio de $\mathbf{R}^n$. Este subespacio W a menudo se denota con $V^\perp$ y se conoce como V **perp**, o también como **complemento ortogonal** de V.

III, §2. Combinaciones lineales

Sea V un espacio vectorial y sean $v_1, \ldots, v_n$ elementos de V. Diremos que $v_1, \ldots, v_n$ **generan** V si, dado cualquier elemento $v \in V$, existen números $x_1, \ldots, x_n$ tales que

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Ejemplo 1. Sea $E_1, \ldots, E_n$ el vector unitario en $\mathbf{R}^n$ de manera que E_i tiene la componente 1 en el lugar i y la componente 0 en todos los demás lugares. Entonces $E_1, \ldots, E_n$ generan $\mathbf{R}^n$. Prueba: dado $X = (x_1, \ldots, x_n) \in \mathbf{R}^n$, entonces

$$X = \sum_{i=1}^{n} x_i E_i,$$

de manera que existen números que satisfacen la condición de la definición.

Sea V un espacio vectorial arbitrario y sean $v_1, \ldots, v_n$ elementos de V. Sean $x_1, \ldots, x_n$ números. A una expresión del tipo

$$x_1 v_1 + \cdots + x_n v_n$$

se le conoce como **combinación lineal** de $v_1, \ldots, v_n$. A los números $x_1, \ldots, x_n$ se les llama entonces **coeficientes** de la combinación lineal.

El conjunto de todas las combinaciones lineales de $v_1, \ldots, v_n$ es un subespacio de V.

Demostración. Sea W el conjunto de todas las combinaciones lineales mencionadas. Sean $y_1, \ldots, y_n$ números. Entonces

$$(x_1 v_1 + \cdots + x_n v_n) + (y_1 v_1 + \cdots + y_n v_n)$$
$$= (x_1 + y_1) v_1 + \cdots + (x_n + y_n) v_n.$$

Así, la suma de dos elementos de W de nuevo es un elemento de W, esto es, una combinación lineal de $v_1, \ldots, v_n$. Además, si c es un número, entonces

$$c(x_1 v_1 + \cdots + x_n v_n) = c x_1 v_1 + \cdots + c x_n v_n$$

es una combinación lineal de $v_1, \ldots, v_n$, y, en consecuencia, es un elemento de W. Por último,

$$O = 0 v_1 + \cdots + 0 v_n$$

es un elemento de W. Esto prueba que W es un subespacio de V.

Se dice que el subespacio W que consiste en todas las combinaciones lineales de $v_1, \ldots, v_n$ es el subespacio generado por $v_1, \ldots, v_n$.

Ejemplo 2. Sea v_1 un elemento no nulo de un espacio vectorial V, y sea w cualquier elemento de V. Al conjunto de elementos

$$w + t v_1, \qquad \text{donde} \qquad t \in \mathbf{R},$$

se le conoce como la **recta que pasa por w en la dirección de v_1**. Ya hemos trabajado con dichas rectas en el Capítulo I, §5. Si $w = O$, entonces la recta que consiste en todos los múltiplos escalares $t v_1$, donde $t \in \mathbf{R}$, es un subespacio generado por v_1.

Sean v_1 y v_2 elementos de un espacio vectorial V y supongamos que ninguno de ellos es un múltiplo escalar del otro. El subespacio generado por v_1 y v_2 se

conoce como el **plano** generado por v_1 y v_2. Consiste en todas las combinaciones lineales

$$t_1v_1 + t_2v_2, \qquad \text{donde} \qquad t_1 \text{ y } t_2 \text{ son números arbitrarios.}$$

Este plano pasa por el origen, tal como se aprecia al poner $t_1 = t_2 = 0$.

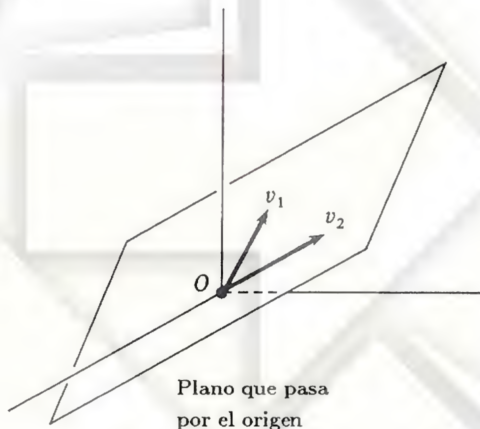


Figura 2

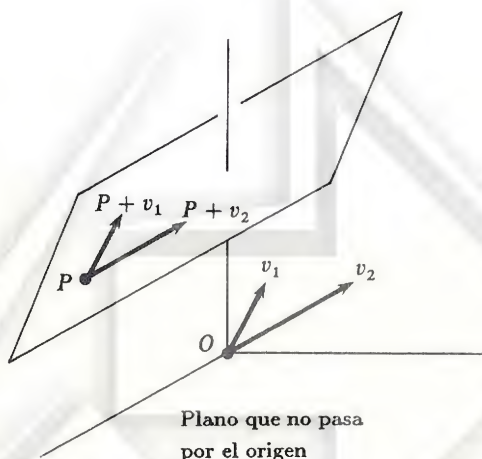
Obtenemos la noción más general de un plano mediante la siguiente operación. Sea S un subconjunto arbitrario de V. Sea P un elemento de V. Si sumamos P a todos los elementos de S, entonces obtenemos lo que se conoce como **traslación** de S determinada por P. Consiste en todos los elementos $P + v$, donde v está en S.

Ejemplo 3. Sean v_1 y v_2 elementos de un espacio vectorial V tales que ninguno de ellos es múltiplo escalar del otro. Sea P un elemento de V. Definimos el **plano que pasa por** P, **paralelo a** v_1, v_2, como el conjunto de todos los elementos

$$P + t_1v_1 + t_2v_2,$$

donde t_1 y t_2 son números arbitrarios. Esta noción de plano es análoga, con dos elementos v_1, v_2, a la noción de recta parametrizada considerada en el Capítulo I.

Advertencia. Usualmente dicho plano no pasa por el origen, como se aprecia en la figura 3. Por tanto, ese plano **no** es un subespacio de V. Sin embargo, si consideramos $P = O$, entonces el plano es un subespacio.



Plano que no pasa por el origen

Figura 3

Algunas veces es interesante restringir los coeficientes de una combinación lineal. En seguida damos algunos ejemplos.

Ejemplo 4. Sea V un espacio vectorial y sean v y u elementos de V. Definimos el **segmento de recta** comprendido entre v y $v+u$ como el conjunto de todos los puntos

$$v + tu, \qquad 0 \leq t \leq 1.$$

En la siguiente figura aparece ilustrado este segmento de recta.

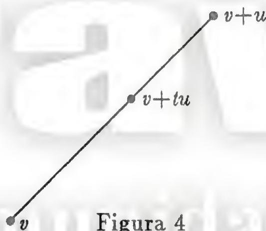


Figura 4

Por ejemplo, si $t = \frac{1}{2}$, entonces $v + \frac{1}{2}u$ es el punto medio entre v y $v + u$. En forma análoga, si $t = \frac{1}{3}$, $v + \frac{1}{3}u$ es el punto que se encuentra a un tercio de la distancia que hay entre v y $v + u$ (Fig. 5).

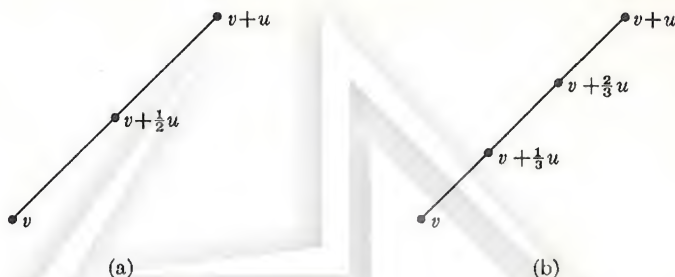


Figura 5

Si v y w son elementos de V, sea $u = w - v$. Entonces el **segmento de recta comprendido entre** v y w es el conjunto de todos los puntos $v + tu$, o

$$v + t(w - v), \qquad 0 \leq t \leq 1.$$

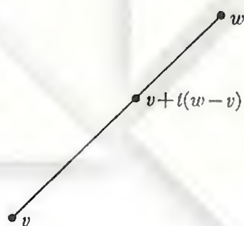


Figura 6

Observe que podemos volver a escribir la expresión para estos puntos en la forma siguiente:

(1) $$(1 - t)v + tw, \qquad 0 \leq t \leq 1,$$

y al hacer $s = 1 - t$, $t = 1 - s$, también la podemos escribir como

$$sv + (1 - s)w, \qquad 0 \leq s \leq 1.$$

Por último podemos escribir los puntos de nuestro segmento de recta en la siguiente forma:

(2) $$t_1 v + t_2 w \quad \text{donde} \quad t_1, t_2 \geq 0 \quad \text{y} \quad t_1 + t_2 = 1.$$

Ciertamente, al hacer $t = t_2$ vemos que todo punto que se puede escribir en la forma (2) satisface (1). Recíprocamente, hacemos $t_1 = 1 - t$ y $t_2 = t$ y vemos que todo punto de la forma (1) se puede escribir en la forma (2).

Ejemplo 5. Sean v y w elementos de un espacio vectorial V. Supongamos que ninguno es múltiplo escalar del otro. Definamos el **paralegramo generado** por v y w como el conjunto de todos los puntos

$$t_1 v + t_2 w, \qquad 0 \leq t_i \leq 1 \qquad \text{para} \qquad i = 1, 2.$$

Esta definición está claramente justificada, puesto que t_1v es un punto del segmento comprendido entre O y v (Fig. 7) y t_2w es un punto del segmento comprendido entre O y w. Para todos los valores de t_1 y t_2 que varían en forma independiente entre 0 y 1, vemos geométricamente que t_1v+t_2w describe todos los puntos del paralelogramo.

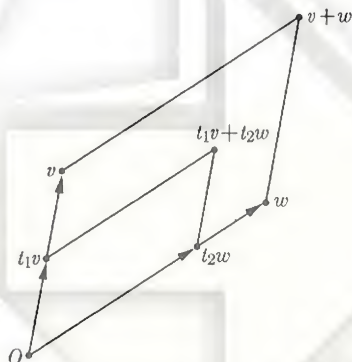


Figura 7

Al considerar la traslación del paralelogramo que se acaba de describir obtenemos el paralelogramo más general (Fig. 8). Así, si u es un elemento de V, la traslación mediante u del paralelogramo generado por v y w consiste en todos los puntos

$$u+t_1v+t_2w, \qquad 0 \leq t_i \leq 1 \qquad \text{para} \qquad i=1,2.$$

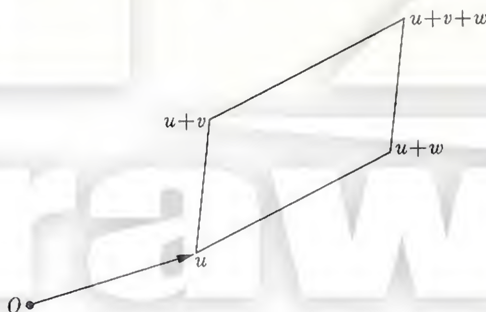


Figura 8

En forma similar, en dimensiones superiores, sean v_1, v_2 y v_3 elementos de un espacio vectorial V. Definamos la **caja generada por estos elementos** como el conjunto de combinaciones lineales

$$t_1v_1+t_2v_2+t_3v_3 \qquad \text{donde} \qquad 0 \leq t_i \leq 1.$$

Hagamos un dibujo en el que v_1, v_2 y v_3 se encuentren en posición general:

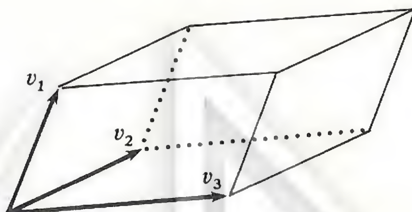


Figura 9

Puede haber casos degenerados que, un poco más adelante, nos conducirán a la noción de dependencia lineal.

Ejercicios III, §2

1. Sean $A_1, \ldots, A_r$ generadores de un subespacio V de $\mathbf{R}^n$. Sea W el conjunto de todos los elementos de $\mathbf{R}^n$ que son perpendiculares a $A_1, \ldots, A_r$. Muestre que los vectores de W son perpendiculares a todo elemento de V.
2. Dibuje el paralelogramo generado por los vectores $(1,2)$ y $(-1,1)$ en $\mathbf{R}^2$.
3. Dibuje el paralelogramo generado por los vectores $(2,-1)$ y $(1,3)$ en $\mathbf{R}^2$.

III, §3. Conjuntos convexos

Sea S un subconjunto de un espacio vectorial V. Diremos que S es **convexo** si, dados los puntos P y Q en S, el segmento de recta comprendido entre P y Q está contenido en S. En la figura 10, el conjunto de la izquierda es convexo. El conjunto de la derecha no es convexo porque el segmento de recta comprendido entre P y Q no está completamente contenido en S.



Figura 10

Recordemos que el segmento de recta comprendido entre P y Q consiste en todos los puntos

$$(1-t)P + tQ \qquad \text{donde} \qquad 0 \le t \le 1.$$

Esto nos da un criterio simple para determinar si un conjunto es convexo o no lo es.

Ejemplo 1. Sea S el paralelogramo generado por dos vectores v_1 y v_2, de manera que S es el conjunto de combinaciones lineales

$$t_1v_1 + t_2v_2 \qquad \text{donde} \qquad 0 \leq t_i \leq 1.$$

Deseamos probar que S es convexo. Sean

$$P = t_1v_1 + t_2v_2 \qquad \text{y} \qquad Q = s_1v_1 + s_2v_2$$

puntos de S. Entonces

$$\begin{aligned}(1-t)P + tQ &= (1-t)(t_1v_1 + t_2v_2) + t(s_1v_1 + s_2v_2)\\ &= (1-t)t_1v_1 + (1-t)t_2v_2 + ts_1v_1 + ts_2v_2\\ &= r_1v_1 + r_2v_2,\end{aligned}$$

donde

$$r_1 = (1-t)t_1 + ts_1 \qquad \text{y} \qquad r_2 = (1-t)t_2 + ts_2.$$

Además tenemos que

$$0 \leq (1-t)t_1 + ts_1 \leq (1-t) + t = 1$$

y

$$0 \leq (1-t)t_2 + ts_2 \leq (1-t) + t = 1.$$

En consecuencia,

$$(1-t)P + tQ = r_1v_1 + r_2v_2 \qquad \text{donde} \qquad 0 \leq r_i \leq 1.$$

Esto prueba que $(1-t)P+tQ$ está en el paralelogramo, el cual, por consiguiente, es convexo.

Ejemplo 2. Semiplanos. Considere una ecuación lineal como la siguiente:

$$2x - 3y = 6.$$

Ésta es la ecuación de una recta como se muestra en la figura 11.

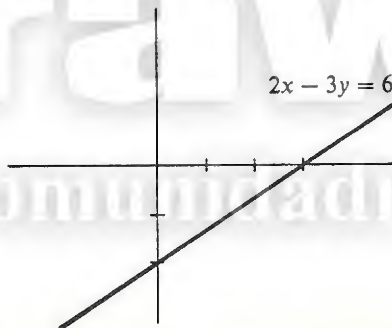


Figura 11

Las desigualdades

$$2x - 3y \leq 6 \qquad \text{y} \qquad 2x - 3y \geq 6$$

determinan dos semiplanos; uno de ellos se encuentra por debajo de la recta y el otro por arriba, como se muestra en la figura 12.

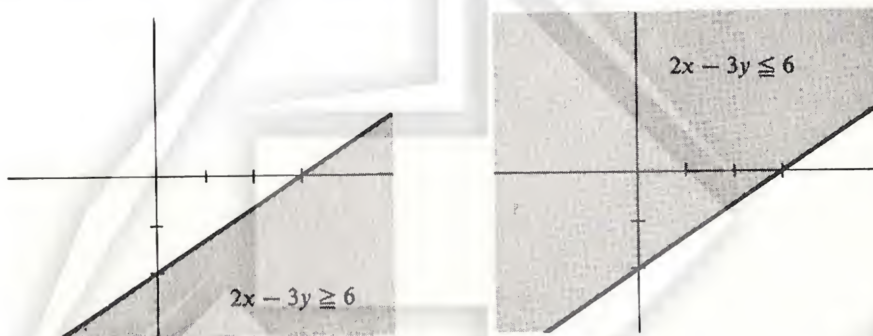


Figura 12

Sea $A = (2, -3)$. Podemos —y deberíamos— escribir las ecuaciones lineales en la forma

$$A \cdot X \geq 6 \qquad \text{y} \qquad A \cdot X \leq 6,$$

donde $X = (x, y)$. Como ejercicio 2, pruebe que todo semiplano es convexo. Esto es intuitivamente claro a partir de la figura, al menos en $\mathbf{R}^2$, aunque su demostración deberá ser válida para la situación análoga en $\mathbf{R}^n$.

Teorema 3.1. *Sean $P_1, \ldots, P_n$ puntos de un espacio vectorial V. Sea S el conjunto de todas las combinaciones lineales*

$$t_1 P_1 + \cdots + t_n P_n$$

donde $0 \leq t_i$ y $t_1 + \cdots + t_n = 1$. Entonces S es convexo.

Demostración. Sean

$$P = t_1 P_1 + \cdots + t_n P_n$$

y

$$Q = s_1 P_1 + \cdots + s_n P_n$$

donde $0 \leq t_i$, $0 \leq s_i$, y

$$t_1 + \cdots + t_n = 1,$$
$$s_1 + \cdots + s_n = 1.$$

Sea $0 \leq t \leq 1$. Entonces:

$$\begin{aligned}(1-t)P + tQ &= (1-t)t_1 P_1 + \cdots + (1-t)t_n P_n \\ &\quad + t s_1 P_1 + \cdots + t s_n P_n \\ &= [(1-t)t_1 + t s_1] P_1 + \cdots + [(1-t)t_n + t s_n] P_n.\end{aligned}$$

Tenemos que $0 \leq (1-t)t_i + ts_i$ para todo i, y

$$\begin{aligned}(1-t)t_1 + ts_1 + \cdots + (1-t)t_n + ts_n \\ &= (1-t)(t_1 + \cdots + t_n) + t(s_1 + \cdots + s_n) \\ &= (1-t) + t \\ &= 1.\end{aligned}$$

Esto prueba nuestro teorema.

En el siguiente teorema probaremos que el conjunto de todas las combinaciones lineales

$$t_1P_1 + \cdots + t_nP_n \qquad \text{donde} \qquad 0 \leq t_i \qquad \text{y} \qquad t_1 + \cdots + t_n = 1$$

es el mínimo conjunto convexo que contiene a $P_1, \ldots, P_n$. Por ejemplo, suponga que P_1, P_2 y P_3 son tres puntos del plano que no están alineados. Geométricamente es claro entonces que el mínimo conjunto convexo que contiene a estos tres puntos es el triángulo que tiene estos puntos como vértices.

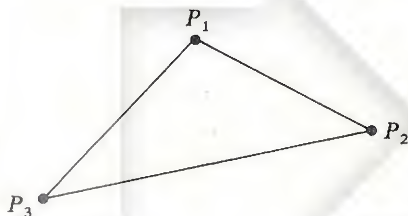


Figura 13

Por ello es natural considerar como definición de triángulo la siguiente propiedad, válida en cualquier espacio vectorial.

Sean P_1, P_2 y P_3 tres puntos de un espacio vectorial V, que no están alineados. Entonces el **triángulo generado** por estos puntos es el conjunto de todas las combinaciones

$$t_1P_1 + t_2P_2 + t_3P_3 \qquad \text{donde} \qquad 0 \leq t_i \qquad \text{y} \qquad t_1 + t_2 + t_3 = 1.$$

Cuando trabajamos con más de tres puntos, el conjunto de combinaciones lineales, como en el teorema 3.1, tiene el aspecto de la siguiente figura.

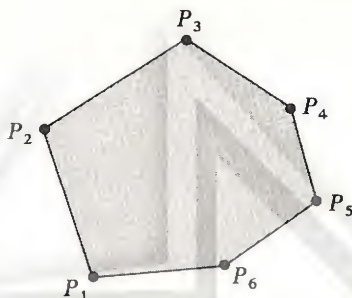


Figura 14

Diremos que el conjunto convexo del teorema 2.1 es el conjunto convexo **generado** por $P_1, \dots, P_n$. Aunque no lo necesitaremos, el siguiente resultado muestra que este conjunto convexo es el mínimo conjunto convexo que contiene todos los puntos $P_1, \dots, P_n$. El lector deberá omitir la prueba si no puede manejar el argumento por inducción.

Teorema 3.2. *Sean $P_1, \dots, P_n$ puntos de un espacio vectorial V. Cualquier conjunto convexo que contiene a $P_1, \dots, P_n$ también contiene todas las combinaciones lineales*

$$t_1P_1 + \cdots + t_nP_n$$

donde $0 \le t_i$ para todo i y $t_1 + \cdots + t_n = 1$.

Demostración. Se hará la prueba por inducción. Si $n = 1$, entonces $t_1 = 1$ y nuestro aserto es obvio. Supongamos que el teorema está probado para algún entero $n - 1 \ge 1$. Lo probaremos para n. Sean $t_1, \dots, t_n$ números que satisfacen las condiciones del teorema. Sea S' un conjunto convexo que contiene a $P_1, \dots, P_n$. Debemos demostrar que S' contiene todas las combinaciones lineales

$$t_1P_1 + \cdots + t_nP_n.$$

Si $t_n = 1$, entonces nuestro aserto es trivial porque $t_1 = \cdots = t_{n-1} = 0$. Supongamos que $t_n \ne 1$. Entonces la combinación lineal $t_1P_1 + \cdots + t_nP_n$ es igual a

$$(1 - t_n)\left(\frac{t_1}{1 - t_n}P_1 + \cdots + \frac{t_n - 1}{1 - t_n}P_n - 1\right) + t_nP_n.$$

Sea

$$S_i = \frac{t_i}{1 - t_i} \qquad \text{para} \qquad i = 1, \dots, n - 1.$$

Entonces $s_i \ge 0$ y $s_1 + \cdots + s_{n-1} = 1$, de manera que, por inducción, concluimos que el punto

$$Q = s_1P_1 + \cdots + s_{n-1}P_{n-1}$$

está en S'. Pero entonces

$$(1 - t_n)Q + t_nP_n = t_1P_1 + \cdots + t_nP_n$$

está en S', por definición de conjunto convexo, tal como se quería probar.

Ejercicios III, §3

1. Sea S el paralelogramo que consiste en todas las combinaciones lineales $t_1v_1 + t_2v_2$, donde $0 \leq t_1 \leq 1$ y $0 \leq t_2 \leq 1$. Pruebe que S es convexo.
2. Sea A un vector no nulo de $\mathbf{R}^n$ y sea c un número fijo. Muestre que el conjunto de todos los elementos X de $\mathbf{R}^n$ tales que $A \cdot X \geq c$ es convexo.
3. Sea S un conjunto convexo de un espacio vectorial. Si c es un número, denote con cS el conjunto de todos los elementos cv, donde v está en S. Demuestre que cS es convexo.
4. Sean S_1 y S_2 conjuntos convexos. Demuestre que la intersección $S_1 \cap S_2$ es convexa.
5. Sea S un conjunto convexo en un espacio vectorial V. Sea w un elemento arbitrario de V. Sea $w + S$ el conjunto de todos los elementos $w + v$, con v en S. Demuestre que $w + S$ es convexo.

III, §4. Independencia lineal

Sea V un espacio vectorial y sean $v_1, \ldots, v_n$ elementos de V. Diremos que $v_1, \ldots, v_n$ son **linealmente dependientes** si existen números $a_1, \ldots, a_n$, no todos iguales a 0, tales que

$$a_1v_1 + \cdots + a_nv_n = O.$$

Si no existen tales números, entonces decimos que $v_1, \ldots, v_n$ son **linealmente independientes**. En otras palabras, los vectores $v_1, \ldots, v_n$ son linealmente independientes si, y sólo si, se satisface la siguiente condición:

Sean $a_1, \ldots, a_n$ números tales que

$$a_1v_1 + \cdots + a_nv_n = O;$$

entonces $a_i = 0$ para todo $i = 1, \ldots, n$.

Ejemplo 1. Sea $V = \mathbf{R}^n$ y considere los vectores

$$E_1 = (1, 0, \ldots, 0)$$
$$\vdots$$
$$E_n = (0, 0, \ldots, 1).$$

Entonces $E_1, \ldots, E_n$ son linealmente independientes. En efecto, sean $a_1, \ldots, a_n$ números tales que $a_1E_1 + \cdots + a_nE_n = O$. Como

$$a_1E_1 + \cdots + a_nE_n = (a_1, \ldots, a_n),$$

se infiere que todo $a_i = 0$.

Ejemplo 2. Muestre que los vectores $(1,1)$ y $(-3,2)$ son linealmente independientes.

Sean a y b dos números tales que

$$a(1,1)+b(-3,2)=O.$$

Al escribir esta ecuación en términos de componentes, encontramos que

$$a-3b=0, \qquad a+2b=0.$$

Éste es un sistema de dos ecuaciones que resolvemos para a y b. Al restar la segunda ecuación de la primera obtenemos $-5b=0$, por lo que $b=0$. Sustituyendo en cualquier ecuación emcontramos que $a=0$. En consecuencia, a y b son iguales a 0 y así nuestros vectores son linealmente independientes.

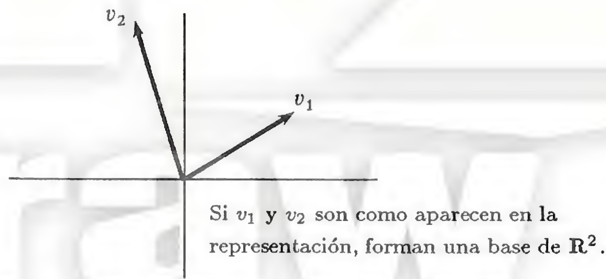
Si los elementos $v_1,\ldots,v_n$ de V generan V y además son linealmente independientes, entonces se dice que $\{v_1,\ldots,v_n\}$ es una **base** de V. También decimos que los elementos $v_1,\ldots,v_n$ **constituyen** o **forman** una base de V.

Ejemplo 3. Los vectores $E_1,\ldots,E_n$ del Ejemplo 1 forman una base de $\mathbf{R}^n$. Para probar esto tenemos que comprobar que son linealmente independientes, lo cual ya se hizo en el Ejemplo 1, y que generan $\mathbf{R}^n$. Dado un elemento $A=(a_1,\ldots,a_n)$ de $\mathbf{R}^n$, podemos expresar A como una combinación lineal

$$A=a_1E_1+\cdots+a_nE_n,$$

de manera que, por definición, $E_1,\ldots,E_n$ generan $\mathbf{R}^n$, por lo que forman una base.

Sin embargo, hay muchas bases más. Consideremos $n=2$. Hallamos que cualesquiera dos vectores que no sean paralelos forman una base de $\mathbf{R}^2$. Primero consideremos un ejemplo.



Si v_1 y v_2 son como aparecen en la representación, forman una base de $\mathbf{R}^2$.

Figura 15

Ejemplo 4. Muestre que los vectores $(1,1)$ y $(-1,2)$ forman una base de $\mathbf{R}^2$.

Tenemos que demostrar que son linealmente independientes y que generan $\mathbf{R}^2$. Para probar la independencia lineal, suponga que a y b son números tales que

$$a(1,1)+b(-1,2)=(0,0).$$

Entonces

$$a - b = 0, \qquad a + 2b = 0.$$

Al restar la primera ecuación de la segunda se obtiene $3b = 0$, de manera que $b = 0$. Pero entonces, de la primera ecuación, $a = 0$, lo que prueba que nuestros vectores son linealmente independientes.

Luego, debemos mostrar que $(1,1)$ y $(-1,2)$ generan $\mathbf{R}^2$. Sea (s,t) un elemento arbitrario de $\mathbf{R}^2$. Tenemos que mostrar que existen números x y y tales que

$$x(1,1) + y(-1,2) = (s,t).$$

En otras palabras, debemos resolver el sistema de ecuaciones

$$\begin{aligned} x - \ \ y &= s, \\ x + 2y &= t. \end{aligned}$$

De nuevo reste la primera ecuación de la segunda. Encontramos que

$$3y = t - s,$$

y en consecuencia,

$$y = \frac{t - s}{3},$$

y por último,

$$x = y + s = \frac{t - s}{3} + s.$$

Esto prueba que $(1,1)$ y $(-1,2)$ generan $\mathbf{R}^2$, y concluye la prueba de que forman una base de $\mathbf{R}^2$.

El resultado general para $\mathbf{R}^2$ se expresa en el siguiente teorema.

Teorema 4.1. *Sean (a,b) y (c,d) dos vectores de $\mathbf{R}^2$.*

(i) *Son linealmente dependientes si, y sólo si, $ad - bc = 0$.*
(ii) *Si son linealmente independientes, entonces forman una base de $\mathbf{R}^2$.*

Demostración. Primero efectúela como ejercicio (vea el Ejercicio 4). Si el lector no puede hacerlo, encontrará la prueba en la sección de respuestas. Se asemeja mucho al procedimiento del Ejemplo 4.

Sea V un espacio vectorial y sea $\{v_1, \ldots, v_n\}$ una base de V. Los elementos de V se pueden representar mediante n-tuplas con respecto a esta base, de la manera siguiente. Si un elemento v de V se escribe como combinación lineal

$$v = x_1 v_1 + \cdots + x_n v_n$$

de los elementos de la base, entonces se conoce a $(x_1, \ldots, x_n)$ como **coordenadas** de v con respecto a nuestra base y a x_i se le conoce como la i-ésima coordenada. Las coordenadas con respecto a la base usual $E_1, \ldots, E_n$ de $\mathbf{R}^n$ simplemente son las coordenadas tal como se definieron en el Capítulo I, §1.

El siguiente teorema muestra que sólo puede haber un conjunto de coordenadas para un vector dado.

Teorema 4.2. *Sea V un espacio vectorial. Sean $v_1, \ldots, v_n$ elementos linealmente independientes de V. Sean $x_1, \ldots, x_n$ y $y_1, \ldots, y_n$ números tales que*

$$x_1v_1 + \cdots + x_nv_n = y_1v_1 + \cdots + y_nv_n.$$

Entonces debemos tener que $x_i = y_i$ para todo $i = 1, \ldots, n$.

Demostración. Restemos el miembro derecho de la igualdad, del izquierdo. Obtenemos

$$x_1v_1 - y_1v_1 + \cdots + x_nv_n - y_nv_n = O.$$

También podemos escribir esta relación en la forma

$$(x_1 - y_1)v_1 + \cdots + (x_n - y_n)v_n = O.$$

Por definición, debemos tener $x_i - y_i = 0$ para todo $i = 1, \ldots, n$, con lo que se ha probado nuestro aserto.

El teorema expresa el hecho de que, cuando un elemento se escribe como combinación lineal de $v_1, \ldots, v_n$, entonces sus coeficientes $x_1, \ldots, x_n$ están determinados en forma única. Esto es cierto sólo cuando $v_1, \ldots, v_n$ son linealmente independientes.

Ejemplo 5. Encuentre las coordenadas de $(1,0)$ con respecto a los dos vectores $(1,1)$ y $(-1,2)$.

Debemos encontrar números a y b tales que

$$a(1,1) + b(-1,2) = (1,0).$$

Al escribir esta ecuación en términos de coordenadas, encontramos

$$a - b = 1, \qquad a + 2b = 0.$$

Al resolver para a y b de la manera usual se obtiene $b = -\frac{1}{3}$ y $a = \frac{2}{3}$. Por tanto las coordenadas de $(1,0)$ con respecto a $(1,1)$ y $(-1,2)$ son $(\frac{2}{3}, -\frac{1}{3})$.

Ejemplo 6. Las dos funciones e^t y e^{2t} son linealmente independientes. Para probar esto, suponga que existen números a y b tales que

$$ae^t + be^{2t} = 0$$

(para todos los valores de t). Derive esta relación. Obtenemos

$$ae^t + 2be^{2t} = 0.$$

Restemos la primera relación de la segunda. Obtenemos $be^t = 0$ y, por tanto, $b = 0$. A partir de la primera relación, se infiere que $ae^t = 0$ y en consecuencia, $a = 0$, por lo que e^t y e^{2t} son linealmente independientes.

Ejemplo 7. Sea V el espacio vectorial de todas las funciones de una variable t. Sean $f_1, \ldots, f_n$, n funciones. Decir que son linealmente dependientes es decir que existen n números $a_1, \ldots, a_n$, no todos iguales a 0, tales que

$$a_1f_1(t) + \cdots + a_nf_n(t) = 0$$

para *todos* los valores de t.

Advertencia. Queremos hacer énfasis en que la dependencia lineal para funciones significa que se cumple la relación anterior para todos los valores de t. Por ejemplo, considere la relación

$$a \operatorname{sen} t + b \cos t = 0,$$

donde a y b son dos números fijos que no son nulos simultáneamente. Puede haber *algunos* valores de t para los cuales se satisfaga la ecuación anterior. Por ejemplo, si $a \neq 0$, entonces podemos resolver

$$\frac{\operatorname{sen} t}{\cos t} = -\frac{b}{a},$$

o, en otras palabras, $\tan t = b/a$ para obtener al menos una solución. Sin embargo, la relación anterior no se puede cumplir para *todos* los valores de t y, en consecuencia, $\operatorname{sen} t$ y $\cos t$ son linealmente independientes como funciones.

Ejemplo 8. Sea V el espacio vectorial de funciones generado por las dos funciones e^t y e^{2t}. Entonces las coordenadas de la función

$$3e^t + 5e^{2t}$$

con respecto a la base $\{e^t, e^{2t}\}$ son $(3, 5)$.

Existe otra forma conveniente para expresar la independencia lineal al trabajar con dos vectores v y w.

Teorema 4.3. *Sean v y w elementos de un espacio vectorial V. Son linealmente dependientes si, y sólo si, uno de ellos es un múltiplo escalar del otro, esto es, existe un número $c \neq 0$ tal que tenemos $v = cw$ o bien $w = cv$.*

Demostración. Se deja como ejercicio; refiérase al Ejercicio 5.

En vista de este teorema, la condición impuesta en varios ejemplos de la sección anterior se podría formular en términos de dos vectores que sean linealmente independientes.

Ejercicios III, §4

1. Demuestre que los siguientes vectores son linealmente independientes.
   (a) $(1, 1, 1)$ y $(0, 1, -2)$ (b) $(1, 0)$ y $(1, 1)$
   (c) $(-1, 1, 0)$ y $(0, 1, 2)$ (d) $(2, -1)$ y $(1, 0)$
   (e) $(\pi, 0)$ y $(0, 1)$ (f) $(1, 2)$ y $(1, 3)$
   (g) $(1, 1, 0), (1, 1, 1),$ y $(0, 1, -1)$ (h) $(0, 1, 1,), (0, 2, 1)$ y $(1, 5, 3)$

2. Exprese el vector X dado, como combinación lineal de los vectores A y B dados y encuentre las coordenadas de X con respecto a A y B.
   (a) $X = (1, 0)$, $A = (1, 1)$, $B = (0, 1)$
   (b) $X = (2, 1)$, $A = (1, -1)$, $B = (1, 1)$
   (c) $X = (1, 1)$, $A = (2, 1)$, $B = (-1, 0)$
   (d) $X = (4, 3)$, $A = (2, 1,)$, $B = (-1, 0)$

3. Encuentre las coordenadas del vector X con respecto a los vectores A, B y C.
   (a) $X = (1,0,0)$, $A = (1,1,1)$, $B = (-1,1,0)$, $C = (1,0,-1)$
   (b) $X = (1,1,1)$, $A = (0,1,-1)$, $B = (1,1,0)$, $C = (1,0,2)$
   (c) $X = (0,0,1)$, $A = (1,1,1)$, $B = (-1,1,0)$, $C = (1,0,-1)$

4. Sean (a, b) y (c, d) dos vectores de $\mathbf{R}^2$.
   (i) Si $ad - bc \neq 0$, demuestre que son linealmente independientes.
   (ii) Si son linealmente independientes, muestre que $ad - bc \neq 0$.
   (iii) Si $ad - bc \neq 0$, entonces demuestre que forman una base de $\mathbf{R}^2$.

5. (a) Sean v y w elementos de un espacio vectorial. Si v y w son linealmente dependientes, muestre que existe un número c tal que $w = cv$, o bien, $v = cw$.
   (b) Recíprocamente, sean v y w elementos de un espacio vectorial y suponga que existe un número c tal que $w = cv$. Demuestre que v y w son linealmente dependientes.

6. Sean $A_1, \dots, A_r$ vectores de $\mathbf{R}^n$ y suponga que son mutuamente perpendiculares; en otras palabras, $A_i \perp A_j$ si $i \neq j$. Suponga además que ninguno de ellos es O. Pruebe que son linealmente independientes.

7. Considere el espacio vectorial de todas las funciones de una variable t. Muestre que las siguientes parejas de funciones son lineamente independientes.
   (a) $1, t$ (b) t, t^2 (c) t, t^4 (d) e^t, t (e) te^t, e^2t (f) $\operatorname{sen} t, \cos t$
   (g) $t, \operatorname{sen} t$ (h) $\operatorname{sen} t$, $\operatorname{sen} 2t$ (i) $\cos t$, $\cos 3t$

8. Considere el espacio vectorial de las funciones definidas para $t > 0$. Demuestre que las siguientes parejas de funciones son linealmente independientes.
   (a) $t, 1/t$ (b) $e^t, \log t$

9. ¿Cuáles son las coordenadas de la función $3 \operatorname{sen} t + 5 \cos t = f(t)$ con respecto a la base $\{\operatorname{sen} t, \cos t\}$?

10. Sea D la derivada d/dt. Sea $f(t)$ tal como aparece en el Ejercicio 9. ¿Cuáles son las coordenadas de la función $Df(t)$ con respecto a la base que aparece en el Ejercicio 9?

En cada uno de los siguientes casos, exhiba una base para el espacio indicado y pruebe que es una base.

11. El espacio de las matrices de 2×2.
12. El espacio de las matrices de $m \times n$.
13. El espacio de las matrices de $n \times n$ cuyas componentes son todas iguales a 0, exceptuando, posiblemente, las componentes diagonales.

14. Las matrices superiormente triangulares, esto es, matrices del siguiente tipo:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

15. (a) El espacio de las matrices simétricas de 2×2.
    (b) El espacio de las matrices simétricas de 3×3.
16. El espacio de las matrices simétricas de $n \times n$.

III, §5. Dimensión

Nos planteamos la siguiente pregunta: ¿podemos encontrar tres elementos linealmente independiente en $\mathbf{R}^2$? Por ejemplo, ¿los elementos

$$A = (1,2), \qquad B = (-5,7), \qquad C = (10,4)$$

son linealmente independientes? Si el lector desarrolla las ecuaciones lineales que expresan la relación

$$xA + yB + zC = O,$$

encontrará que puede resolverlas para x, y y z no iguales a 0. A saber, estas ecuaciones son:

$$\begin{aligned} x - 5y + 10z &= 0, \\ 2x + 7y + \ \ 4z &= 0. \end{aligned}$$

Éste es un sistema de dos ecuaciones homogéneas con tres incógnitas y sabemos, por el Teorema 2.1 del Capítulo II, que podemos encontrar una solución no trivial (x, y, z) distinta de cero. En consecuencia, A, B y C son linealmente dependientes.

En un momento veremos que éste es un fenómeno general. En $\mathbf{R}^n$ no podemos encontrar más de n vectores linealmente independientes. Más aún, veremos que cualesquiera n elementos de $\mathbf{R}^n$ linealmente independientes deben generar a $\mathbf{R}^n$ y, en consecuencia, deben formar una base. Por último, veremos también que, si una base de un espacio vectorial tiene n elementos y otra base tiene m elementos, entonces, $m = n$. En pocas palabras, dos bases deben tener el mismo número de elementos. Esta propiedad nos permitirá definir la **dimensión** de un espacio vectorial como el número de elementos de cualquier base. Ahora desarrollaremos estas ideas en forma sistemática.

Teorema 5.1. *Sea V un espacio vectorial y supongamos que $\{v_1, \ldots, v_m\}$ generan V. Sean $w_1, \ldots, w_n$ elementos de V y supongamos que $n > m$. Entonces $w_1, \ldots, w_n$ son linealmente dependientes.*

Demostración. Como $\{v_1, \ldots, v_m\}$ genera V, existen números (a_{ij}) tales que podemos escribir

$$\begin{aligned} w_1 &= a_{11}v_1 + \cdots + a_{m1}v_m \\ &\ \ \vdots \\ w_n &= a_{1n}v_1 + \cdots + a_{mn}v_m. \end{aligned}$$

Si $x_1, \ldots, x_n$ son números, entonces

$$\begin{aligned} &x_1w_1 + \cdots + x_nw_n \\ &\qquad = (x_1a_{11} + \cdots + x_na_{1n})v_1 + \cdots + (x_1a_{m1} + \cdots + x_na_{mn})v_m \end{aligned}$$

(basta sumar los coeficientes de $v_1, \ldots, v_m$ verticalmente hacia abajo). Conforme al Teorema 2.1 del Capítulo II, el sistema de ecuaciones

$$\begin{aligned} x_1a_{11} + \cdots + x_na_{1n} &= 0 \\ &\ \ \vdots \\ x_1a_{m1} + \cdots + x_na_{mn} &= 0 \end{aligned}$$

tiene una solución no trivial, debido a que $n > m$. En vista de la observación anterior, dicha solución $(x_1, \ldots, x_n)$ es tal que

$$x_1 w_1 + \cdots + x_n w_n = 0,$$

como se deseaba.

Teorema 5.2. *Sea V un espacio vectorial y supongamos que una base tiene n elementos y que otra tiene m elementos. Entonces $m = n$.*

Demostración. Apliquemos el Teorema 5.1 a las dos bases. El Teorema 5.1 implica que tanto $n > m$ como $m > n$ son imposibles y, en consecuencia, $m = n$.

Sea V un espacio vectorial que tiene una base consistente en n elementos. Diremos que n es la **dimensión** de V. Si V consta sólo del O, entonces V no tiene base y diremos que V tiene dimensión 0.

Ahora podemos reformular las definiciones de recta y plano en un espacio vectorial arbitrario V. Una **recta que pasa por el origen** simplemente es un subespacio de una dimensión. Un **plano que pasa por el origen** simplemente es un subespacio de dos dimensiones.

Una **recta** arbitraria se obtiene como la traslación de un subespacio de una dimensión. Un **plano** arbitrario se obtiene como la traslación de un subespacio de dos dimensiones. Cuando una base $\{v_1\}$ ha sido seleccionada para un espacio de una dimensión, entonces los puntos de una recta se expresan en la forma usual

$$P + t_1 v_1 \quad \text{con todos los números posibles} \quad t_1.$$

Cuando se ha seleccionado una base $\{v_1, v_2\}$ para un espacio de dos dimensiones, los puntos de un plano se expresan en la forma

$$P + t_1 v_1 + t_2 v_2 \quad \text{con todos los números posibles} \quad t_1, t_2.$$

Sea $\{v_1, \ldots, v_n\}$ un conjunto de elementos de un espacio vectorial V. Sea r un entero positivo $\leq n$. Diremos que $\{v_1, \ldots, v_r\}$ es un subconjunto **maximal** de elementos linealmente independientes si $v_1, \ldots, v_r$ son linealmente independientes y si, además, dado cualquier v_i, donde $i > r$, los elementos $v_1, \ldots, v_r$, v_i son linealmente dependientes.

El siguiente teorema nos da un útil criterio para determinar cuándo un conjunto de elementos de un espacio vectorial es una base.

Teorema 5.3. *Sea $\{v_1, \ldots, v_n\}$ un conjunto de generadores de un espacio vectorial V. Sea $\{v_1, \ldots, v_r\}$ un subconjunto maximal de elementos linealmente independientes. Entonces $\{v_1, \ldots, v_r\}$ es una base de V.*

Demostración. Debemos probar que $v_1, \ldots, v_r$ generan V. Primero probaremos que cada v_i (para $i > r$) es una combinación lineal de $v_1, \ldots, v_r$. Por hipótesis, dado v_i, existen números $x_1, \ldots, x_r$, y, no todos iguales a 0, tales que

$$x_1 v_1 + \cdots + x_r v_r + y v_i = O.$$

Además $y \neq 0$, porque de otro modo tendríamos una relación de dependencia lineal para $v_1, \ldots, v_r$. De manera que podemos despejar v_i, a saber,

$$v_i = \frac{x_1}{-y}v_1 + \cdots + \frac{x_r}{-y}v_r,$$

con lo que se demuestra que v_i es una combinación lineal de $v_1, \ldots, v_r$.

Luego, sea v cualquier elemento de V. Existen números $c_1, \ldots, c_n$ tales que

$$v = c_1v_1 + \cdots + c_nv_n.$$

En esta relación podemos reemplazar cada $v_i(i > r)$ por una combinación lineal de $v_1, \ldots, v_r$. Si hacemos esto y luego agrupamos términos, encontraremos que hemos expresado v como una combinación lineal de $v_1, \ldots, v_r$. Esto prueba que $v_1, \ldots, v_r$ generan V y, por tanto, forman una base de V.

Daremos ahora criterios que nos permitan decir cuándo los elementos de un espacio vectorial constituyen una base.

Sean $v_1, \ldots, v_n$ elementos linealmente independientes de un espacio vectorial V. Diremos que forman un **conjunto maximal de elementos de V linealmente independientes** si, dado cualquier elemento w de V, los elementos $w, v_1, \ldots, v_n$ son linealmente dependientes.

Teorema 5.4. *Sea V un espacio vectorial y sea $\{v_1, \ldots, v_n\}$ un conjunto maximal de elementos de V linealmente independientes. Entonces $\{v_1, \ldots, v_n\}$ es una base de V.*

Demostración. Ahora debemos demostrar que $v_1, \ldots, v_n$ generan V, esto es, que todo elemento de V se puede expresar como una combinación lineal de $v_1, \ldots, v_n$. Sea w un elemento de V. Los elementos w, $v_1, \ldots, v_n$ de V deben ser linealmente dependientes, por hipótesis, de modo que existen números x_0, $x_1, \ldots, x_n$, no todos iguales a 0, tales que

$$x_0w + x_1v_1 + \cdots + x_nv_n = O.$$

No podemos tener $x_0 = 0$, ya que, si ése fuera el caso, obtendríamos una relación de dependencia lineal entre $v_1, \ldots, v_n$. Por consiguiente, podemos despejar w en términos de $v_1, \ldots, v_n$, a saber,

$$w = -\frac{x_1}{x_0}v_1 - \cdots - \frac{x_n}{x_0}v_n.$$

Esto prueba que w es una combinación lineal de $v_1, \ldots, v_n$, y que $\{v_1, \ldots, v_n\}$ es una base.

Teorema 5.5. *Sea V un espacio vectorial de dimensión n y sean $v_1, \ldots, v_n$ elementos de V linealmente independientes. Entonces $v_1, \ldots, v_n$ constituyen una base de V.*

Demostración. Conforme al Teorema 5.1, $\{v_1, \ldots, v_n\}$ es un conjunto maximal de elementos de V linealmente independientes. En consecuencia, es una base por el Teorema 5.4.

Teorema 5.6. *Sea V un espacio vectorial de dimensión n y sea W un subespacio, también de dimensión n. Entonces $W = V$.*

Demostración. Una base para W también debe ser una base para V.

Teorema 5.7. *Sea V un espacio vectorial de dimensión n. Sea r un entero positivo, donde $r < n$, y sean $v_1, \dots, v_r$ elementos de V linealmente independientes. Entonces se pueden encontrar elementos $v_{r+1}, \dots, v_n$ tales que*

$$\{v_1, \dots, v_n\}$$

es una base de V.

Demostración. Como $r < n$, sabemos que $\{v_1, \dots, v_r\}$ no puede formar una base de V y, por tanto, no puede ser un conjunto maximal de elementos de V linealmente independientes. En particular, podemos encontrar v_{r+1} en V tal que

$$v_1, \dots, v_{r+1}$$

son linealmente independientes. Si $r + 1 < n$, entonces podemos repetir el argumento. Así podemos proceder paso a paso (por inducción) hasta que obtenemos n elementos linealmente independientes $\{v_1, \dots, v_n\}$. Éstos deben ser una base, por el Teorema 5.4, y nuestro corolario queda probado.

Teorema 5.8. *Sea V un espacio vectorial que tiene una base que consta de n elementos. Sea W un subespacio que no consta sólo del O. Entonces W tiene una base y la dimensión de W es $\leq n$.*

Demostración. Sea w_1 un elemento no nulo de W. Si $\{w_1\}$ no es un conjunto maximal de elementos de W linealmente independientes, entonces podemos encontrar un elemento w_2 de W tal que w_1 y w_2 son linealmente independientes. Procediendo de esta manera, un elemento cada vez, debe haber un entero $m \leq n$ tal que podemos encontrar elementos $w_1, w_2, \dots, w_m$ linealmente independientes y tales que

$$\{w_1, \dots, w_m\}$$

es un conjunto maximal de elementos de W linealmente independientes (por el Teorema 5.1 no podemos continuar encontrando en forma indefinida elementos linealmente independientes, y el número de tales elementos es, a lo más, n). Si ahora usamos el Teorema 5.4, concluimos que $\{w_1, \dots, w_m\}$ es una base para W.

Ejercicios III, §5

1. ¿Cuál es la dimensión de los siguientes espacios? (Consulte los ejercicios 11 a 16 de la sección anterior):
   (a) Las matrices de 2×2. (b) Las matrices de $m \times n$.

(c) Las matrices de $n \times n$ cuyas componentes son todas iguales a 0, exceptuando posiblemente las de la diagonal.
(d) Las matrices superiormente triangulares de $n \times n$.
(e) Las matrices simétricas de 2×2.
(f) Las matrices simétricas de 3×3.
(g) Las matrices simétricas de $n \times n$.

2. Sea V un subespacio de $\mathbf{R}^2$. ¿Cuáles son las posibles dimensiones de V? Muestre que, si $V \neq \mathbf{R}^2$, entonces o bien $V = \{O\}$, o V es una recta que pasa por el origen.

3. Sea V un subespacio de $\mathbf{R}^3$. ¿Cuáles son las posibles dimensiones de V? Muestre que, si $V \neq \mathbf{R}^3$, entonces, o bien $V = \{O\}$, o V es una recta que pasa por el origen, o V es un plano que pasa por el origen.

III, §6. El rango de una matriz

Sea

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

una matriz de $m \times n$. Las columnas de A generan un espacio vectorial, que es un subespacio de $\mathbf{R}^m$. La dimensión de ese subespacio se llama **rango por columnas** de A. En vista del Teorema 5.4, el rango por columnas es igual al número máximo de columnas linealmente independientes. En forman análoga, los renglones de A generan un subespacio de $\mathbf{R}^n$ y la dimensión de este subespacio se conoce como **rango por renglones**. De nuevo por el Teorema 5.4, el rango por renglones es igual al número máximo de renglones linealmente independientes. Más adelante probaremos que estos dos rangos son iguales entre sí. Daremos dos pruebas; la primera, de esta sección, depende de ciertas operaciones por renglones y por columnas de una matriz. Posteriormente daremos una prueba más geométrica empleando la noción de perpendicularidad.

Definimos el **espacio de renglones** de A como el subespacio generado por los renglones de A. Definimos el **espacio de columnas** de A como el subespacio generado por las columnas.

Considere las siguientes operaciones por renglones de una matriz.

Renglón 1. Sumar un múltiplo escalar de un renglón a otro.

Renglón 2. Intercambiar renglones.

Renglón 3. Multiplicar un renglón por un escalar no nulo.

Éstas se conocen como operaciones por renglones (algunas veces como operaciones elementales por renglones). Tenemos operaciones semejantes para las columnas, las que se denotarán con **Col 1, Col 2** y **Col 3**, respectivamente. Estudiaremos el efecto de estas operaciones sobre los rangos.

Observe primero que cada una de las operaciones anteriores tiene una operación inversa en el sentido de que, efectuando operaciones similares, podemos regresar a la matriz original. Por ejemplo, cambiemos una matriz A sumando c veces el segundo renglón al primero. Obtenemos una nueva matriz B cuyos renglones son

$$B_1 = A_1 + cA_2, A_2, \ldots, A_m.$$

Si ahora sumamos $-cA_2$ al primer renglón de B, obtenemos otra vez A_1. Se puede aplicar un argumento semejante a dos renglones cualesquiera.

Si intercambiamos dos renglones y luego los intercambios de nuevo, volvemos a obtener la matriz original.

Si multiplicamos un renglón por un número $c \neq 0$, entonces, al multiplicar de nuevo por c^{-1}, se obtiene el renglón original.

Teorema 6.1. *Las operaciones por renglones y columnas no cambian el rango por renglones de una matriz, así como tampoco el rango por columnas.*

Demostración. Observemos primero que al intercambiar renglones de una matriz no se afecta el rango por renglones, puesto que el subespacio generado por los renglones es el mismo, sin importar en qué orden consideramos los renglones.

Luego suponga que sumamos un múltiplo escalar de un renglón a otro. Mantendremos la notación que usamos antes del teorema, de manera que los nuevos renglones son

$$B_1 = A_1 + cA_2, A_2, \ldots, A_m.$$

Cualquier combinación lineal de los renglones de B, a saber, cualquier combinación de

$$B_1, A_2, \ldots, A_m$$

también es una combinación lineal de $A_1, A_2, \ldots, A_m$. En consecuencia, el espacio de los renglones de B está contenido en el espacio de los renglones de A. Por tanto, por el Teorema 5.6 tenemos

$$\text{rango por renglones de } B \leq \text{rango por renglones de } A.$$

Como A también se obtiene de B mediante una operación similar, obtenemos la desigualdad inversa

$$\text{rango por renglones de } A \leq \text{rango por renglones de } B.$$

Por tanto, estos dos rangos por renglones son iguales.

Tercero, si multiplicamos un renglón A_i por $c \neq 0$, obtenemos el nuevo renglón cA_i. Pero $A_i = c^{-1}(cA_i)$, de manera que los espacios de renglones de la matriz A y los de la nueva matriz obtenida al multiplicar el renglón por c son iguales. Así, la tercera operación tampoco cambia el rango por renglones.

Podríamos haber dado el argumento anterior para cualquier par de renglones A_i, $A_j (i \neq j)$, de manera que hemos probado que las operaciones por renglones no cambian el rango por renglones.

Ahora probaremos que no cambia el rango por columnas.

Considere de nuevo la matriz obtenida al sumar un múltiplo escalar del segundo renglón al primero:

$$B = \begin{pmatrix} a_{11} + ca_{21} & a_{12} + ca_{22} & \cdots & a_{1n} + ca_{2n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Sean $B^1, \ldots, B^n$ las columnas de esta nueva matriz B. Veremos que la relación de dependencia lineal entre las columnas de B es precisamente la misma que la relación de dependencia lineal entre las columnas de A. En otras palabras:

Un vector $X = (x_1, \ldots, x_n)$ da una relación de dependencia lineal

$$x_1 B^1 + \cdots + x_n B^n = O$$

entre las columnas de B si, y sólo si, X da una relación de dependencia lineal

$$x_1 A^1 + \cdots + x_n A^n = O$$

entre las columnas de A.

Demostración. A partir del Capítulo II, sección §2, sabemos que una relación de dependencia lineal entre las columnas se puede escribir en términos del producto interior con los renglones de la matriz. Así, suponga que tenemos una relación

$$x_1 B^1 + \cdots + x_n B^n = O.$$

Esto es equivalente al hecho de que

$$X \cdot B_i = 0 \qquad \text{para} \qquad i = 1, \ldots, m.$$

Por consiguiente,

$$X \cdot (A_1 + cA_2) = 0, \quad X \cdot A_2 = 0, \quad \ldots, \quad X \cdot A_m = 0.$$

La primera ecuación se puede escribir de la siguiente manera:

$$X \cdot A_1 + cX \cdot A_2 = 0.$$

Como $X \cdot A_2 = 0$, concluimos que $X \cdot A_1 = 0$. En consecuencia, X es perpendicular a los renglones de A y, por tanto, X da una relación lineal entre las columnas de A. El recíproco se prueba de manera semejante.

El enunciado anterior prueba que, si r columnas de B son linealmente independientes, entonces r columnas de A también son linealmente independientes, y recíprocamente. Por consiguiente, A y B tienen el mismo rango por columnas.

Dejamos que el lector verifique que las otras dos operaciones por renglones no cambian los rangos por columnas.

Del mismo modo se prueba que las operaciones por columnas no cambian el rango por renglones. La situación es simétrica entre renglones y columnas. Esto concluye la prueba del teorema.

Teorema 6.2. *Sea A una matriz de rango por renglones igual a r. Mediante una sucesión de operaciones por renglones y por columnas, la matriz se puede transformar en la matriz que tiene por componentes a 1 en la diagonal de los primeros r renglones y columnas y 0 en las demás componentes.*

$$\begin{pmatrix} 1 & 0 & \cdots & & 0 & 0 \\ 0 & 1 & \cdots & & 0 & \vdots \\ \vdots & \vdots & \ddots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}$$

En particular, el rango por renglones es igual al rango por columnas.

Demostración. Suponga que $r \neq 0$, de manera que la matriz no es la matriz nula. Alguna componente no es igual a cero. Después de intercambiar renglones y columnas, de ser necesario, podemos suponer que esta componente está en la esquina superior izquierda, esto es, esta componente es igual a $a_{11} \neq 0$. Ahora descendemos por la primera columna; multipliquemos el primer renglón por a_{21}/a_{11} y restémoslo del segundo renglón. Entonces obtenemos una matriz que tiene 0 en el primer lugar del segundo renglón. Luego multipliquemos el primer renglón por a_{31}/a_{11} y restémoslo del tercer renglón. Entonces nuestra nueva matriz tiene la primera componente igual a 0 en el tercer renglón. Procediendo de la misma manera, podemos transformar la matriz de manera que sea de la forma

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Después, restemos múltiplos apropiados de la primera columna, de la segunda, de la tercera,..., de la n-ésima columna para obtener ceros en el primer renglón. Esto transforma a la matriz en una matriz del siguiente tipo:

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Ahora tenemos una matriz de $(m-1) \times (n-1)$ en la parte inferior de la derecha. Si efectuamos operaciones por renglones y columnas en todas partes excepto en el primer renglón y en la primera columna, entonces, primero, no modificamos la primera componente a_{11}; y segundo, podemos repetir el argumento,

con elobjeto de obtener una matriz de la forma siguiente:

$$\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{m3} & \cdots & a_{mn} \end{pmatrix}.$$

Procediendo paso a paso mediante inducción llegamos a una matriz de la forma

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 & 0 \\ 0 & a_{22} & \cdots & 0 & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & a_{ss} & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

cuyos elementos diagonales $a_{11}, \ldots, a_{ss}$ son $\neq 0$. Dividamos el primer renglón entre a_{11}, el segundo entre a_{22}, etc. Entonces obtendremos una matriz

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Por tanto, tenemos la matriz unitaria de $s \times s$ en la esquina superior de la izquierda y ceros en todo lo demás. Como las operaciones por renglones y por columnas no cambian el rango por renglones o por columnas, se infiere que $r = s$ y también que el rango por renglones es igual al rango por columnas. Esto prueba el teorema.

Puesto que hemos probado que el rango por renglones es igual al rango por columnas, podemos ahora omitir "renglones" o "columnas" y solo hablar del **rango** de una matriz. Así, por definición, el **rango** de una matriz es igual a la dimensión del espacio generado por los renglones.

Observación. Aunque el procedimiento sistemático suministra un método efectivo para encontrar el rango, en la práctica usualmente se buscan atajos para obtener tantos ceros como sea posible al efectuar las operaciones por renglones y por columnas, de manera que en algún punto sea obvio cuál es el rango de la matriz.

Por supuesto, también se puede emplear el simple mecanismo de las ecuaciones lineales para encontrar el rango.

Ejemplo. Encuentre el rango de la matriz

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

Sólo hay dos renglones, de manera que el rango es a lo más igual a 2. Por otro lado, las dos columnas

$$\begin{pmatrix} 2 \\ 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Son linealmente independientes ya que si a y b son números tales que

$$a\begin{pmatrix} 2 \\ 0 \end{pmatrix} + b\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

entonces

$$\begin{aligned} 2a + b &= 0, \\ b &= 0, \end{aligned}$$

de manera que $a = 0$. Por consiguiente, las dos columnas son linealmente independientes y el rango es igual a 2.

Más adelante también veremos que los determinantes dan una forma de computación para determinar cuándo los vectores son linealmente independientes y, por tanto, se pueden usar para determinar el rango.

Ejemplo. Encuentre el rango de la siguiente matriz.

$$\begin{pmatrix} 1 & 2 & -3 \\ 2 & 1 & 0 \\ -2 & -1 & 3 \\ -1 & 4 & -2 \end{pmatrix}.$$

Restemos dos veces la primera columna de la segunda y sumemos 3 veces la primera columna a la tercera. Esto da

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & -3 & 6 \\ -2 & 3 & -3 \\ -1 & 6 & -5 \end{pmatrix}.$$

Sumemos 2 veces la segunda columna a la tercera. Esto da

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & -3 & 0 \\ -2 & 3 & 3 \\ -1 & 6 & 7 \end{pmatrix}.$$

Esta matriz está escrita en la forma escalonada por columnas y se ve de inmediato que los primeros tres renglones o columnas son linealmente independientes. Como sólo hay tres columnas, se infiere que el rango es igual a 3.

Ejercicios III, §6

1. Encuentre el rango de las siguientes matrices.

(a) $\begin{pmatrix} 2 & 1 & 3 \\ 7 & 2 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} -1 & 2 & -2 \\ 3 & 4 & -5 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 2 & 7 \\ 2 & 4 & -1 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & 2 & -3 \\ -1 & -2 & 3 \\ 4 & 8 & -12 \\ 0 & 0 & 0 \end{pmatrix}$

(e) $\begin{pmatrix} 2 & 0 \\ 0 & -5 \end{pmatrix}$ (f) $\begin{pmatrix} -1 & 0 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 7 \end{pmatrix}$

(g) $\begin{pmatrix} 2 & 0 & 0 \\ -5 & 1 & 2 \\ 3 & 8 & -7 \end{pmatrix}$ (h) $\begin{pmatrix} 1 & 2 & -3 \\ -1 & -2 & 3 \\ 4 & 8 & -12 \\ 1 & -1 & 5 \end{pmatrix}$

(i) $\begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 2 & 2 & 1 \\ 3 & 4 & 2 & 3 \end{pmatrix}$

2. Sea A una matriz triangular

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

Suponga que ninguno de los elementos diagonales es igual a 0. ¿Cuál es el rango de A?

3. Sea A una matriz de $m \times n$ y sea B una matriz de $n \times r$, de manera que podemos formar el producto AB.

(a) Demuestre que las columnas de AB son combinaciones lineales de A. Por tanto, pruebe que

$$\text{rango de } AB \leq \text{rango de } A.$$

(b) Pruebe que rango de $AB \leq$ rango de B. [*Sugerencia*: Use el hecho de que rango de $AB =$ rango de ${}^t(AB)$, y que rango de $B =$ rango de ${}^tB$.]

CAPÍTULO IV

Aplicaciones lineales

Primero definiremos el concepto general de aplicación, que a su vez generaliza el concepto de función. Entre las aplicaciones, las lineales son las más importantes. Una buena parte de las matemáticas se dedica a reducir problemas sobre aplicaciones arbitrarias a aplicaciones lineales, ya que éstas resultan interesantes por sí mismas, además de que muchas aplicaciones son lineales. Por otra parte, a menudo es posible aproximar una aplicación arbitraria mediante una lineal, cuyo estudio es mucho más sencillo que el estudio de la aplicación original. Esto se hace en el cálculo de varias variables.

IV, §1. Aplicaciones

Sean S y S' dos conjuntos. Una **aplicación** de S en S' es una asociación tal que a todo elemento de S le asocia un elemento de S'. En lugar de decir que F es una aplicación de S en S', a menudo escribiremos los símbolos $F: S \to S'$. A una aplicación también se le conoce como **transformación**.

Una función es un tipo especial de aplicación, a saber, es una aplicación de un conjunto en el conjunto de números, es decir, en $\mathbf{R}$.

Extendemos a las aplicaciones algunos de los términos que hemos usado para funciones. Por ejemplo, si $T: S \to S'$ es una aplicación, y si u es un elemento de S, entonces denotamos con $T(u)$, o Tu, el elemento de S' asociado a u mediante T. Decimos que $T(u)$ es el **valor** de T en u, o también que es la **imagen** de u bajo T. Los símbolos $T(u)$ se leen como "T de u". Al conjunto de todos los elementos $T(u)$, cuando u varía sobre todos los elementos de S,

se le conoce como la **imagen** de T. Si W es un subconjunto de S, entonces el conjunto de elementos $T(w)$, cuando w varía sobre todos los elementos de W, recibe el nombre de **imagen** de W bajo T y se denota con $T(W)$.

Sea $F: S \to S'$ una aplicación de un conjunto S en un conjunto S'. Si x es un elemento de S, con frecuencia escribimos

$$x \mapsto F(x)$$

con una flecha especial $\mapsto$ para denotar la imagen de x bajo F. Así, por ejemplo, hablaríamos de la aplicación F tal que $F(x) = x^2$ como la aplicación $x \mapsto x^2$.

Ejemplo 1. Para cualquier conjunto S tenemos la aplicación identidad $I: S \to S$. Se define mediante $I(x) = x$ para todo x.

Ejemplo 2. Sean S y S' iguales a $\mathbf{R}$. Sea $f: \mathbf{R} \to \mathbf{R}$ la función $f(x) = x^2$ (es decir, la función cuyo valor en un número x es x^2). Entonces f es una aplicación de $\mathbf{R}$ en $\mathbf{R}$. Su imagen es el conjunto de números ≥ 0.

Ejemplo 3. Sea S el conjunto de números ≥ 0 y sea $S' = \mathbf{R}$. Sea $g: S \to S'$ la función tal que $g(x) = x^{1/2}$. Entonces g es una aplicación de S en $\mathbf{R}$.

Ejemplo 4. Sea S el conjunto de funciones que tienen derivadas de todos los órdenes en el intervalo $0 < t < 1$ y sea $S' = S$. Entonces la derivada $D = d/dt$ es una aplicación de S en S. En realidad, nuestra aplicación D asocia la función $df/dt = Df$ a la función f. Conforme a nuestra terminología, Df es el valor de la aplicación D en f.

Ejemplo 5. Sea S el conjunto $\mathbf{R}^3$, esto es, el conjunto de ternas. Sea $A = (2, 3, -1)$. Sea $L: \mathbf{R}^3 \to \mathbf{R}$ la aplicación cuyo valor en un vector $X = (x, y, z)$ es $A \cdot X$. Entonces $L(X) = A \cdot X$. Si $X(1, 1, -1)$, entonces el valor de L en X es 6.

Al igual que con las funciones, describimos una aplicación al dar sus valores. Así, en lugar de hacer el enunciado que aparece en el Ejemplo 5 para describir la aplicación L, también diremos: sea $L: \mathbf{R}^3 \to \mathbf{R}$ la aplicación $L(X) = A \cdot X$. Esto es un tanto incorrecto, pero es más breve, y usualmente no se presta a confusión. En forma más correcta podemos escribir $X \mapsto L(X)$ o bien $X \mapsto A \cdot X$, con la flecha especial $\mapsto$ para denotar el efecto de la aplicación L en el elemento X.

Ejemplo 6. Sea $F: \mathbf{R}^2 \to \mathbf{R}^2$ la aplicación dada por

$$F(x, y) = (2x, 2y).$$

Describa la imagen bajo F de los puntos que están en el círculo $x^2 + y^2 = 1$.

Sea (x, y) un punto del círculo de radio 1.

Sean $u = 2x$ y $v = 2y$. Entonces u y v satisfacen la relación

$$(u/2)^2 + (v/2)^2 = 1$$

o, en otras palabras,

$$\frac{u^2}{4} + \frac{v^2}{4} = 1.$$

En consecuencia, (u, v) es un punto del círculo de radio 2, por lo cual la imagen bajo F del círculo de radio 1 es un subconjuto del círculo de radio 2. Recíprocamente, dado un punto (u, v) tal que

$$u^2 + v^2 = 4,$$

sean $x = u/2$ y $y = v/2$. Entonces el punto (x, y) satisface la ecuación

$$x^2 + y^2 = 1,$$

y, en consecuencia, es un punto del círculo de radio 1. Además,

$$F(x, y) = (u, v).$$

Por tanto, todo punto del círculo de radio 2 es la imagen de algún punto del círculo de radio 1. Concluimos que la imagen del círculo de radio 1 bajo F es precisamente el círculo de radio 2.

Nota. En general, sean S y S' dos conjuntos. Para probar que $S = S'$, con frecuencia se prueba que S es un subconjunto de S' y que S' es un subconjunto de S. Esto es lo que hicimos en el argumento anterior.

Ejemplo 7. Este ejemplo es particularmente importante en las aplicaciones geométricas. Sea V un espacio vectorial y sea u un elemento fijo de V. Supongamos que

$$T_u: V \to V$$

es la aplicación tal que $T_u(v) = v + u$. A T_u se le conoce como **traslación determinada por** u. Si S es cualquier subconjunto de V, entonces se conoce a $T_u(S)$ como la **traslación de** S **determinada por** u y consta de todos los vectores $v + u$, donde $v \in S$. Solemos denotarla con $S + u$. En la siguiente figura dibujamos un conjunto S y su traslación determinada por un vector u.

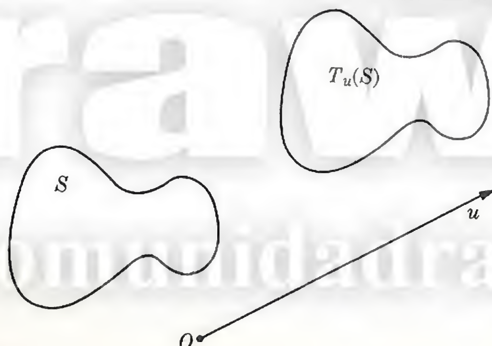


Figura 1

Ejemplo 8. La rotación levógira alrededor del origen determinada por un ángulo θ es una aplicación, que podemos denotar con R_θ. Sea $\theta = \pi/2$. La imagen del punto $(1,0)$ bajo la rotación $R_{\pi/2}$ es el punto $(0,1)$. Podemos escribirlo como

$$R_{\pi/2}(1,0) = (0,1).$$

Ejemplo 9. Sea S un conjunto. Una aplicación de S en $\mathbf{R}$ será denominada **función** y el conjunto de tales funciones se llamará conjunto de funciones definidas en S. Sean f y g dos funciones definidas en S. Podemos definir su suma de la misma manera que para las funciones de números, a saber, $f+g$ es la función cuyo valor en un elemento t de S es $f(t)+g(t)$. También podemos definir el producto de f por un número c. Es la función cuyo valor en t es $cf(t)$. Entonces el conjunto de aplicaciones de S en $\mathbf{R}$ es un espacio vectorial.

Ejemplo 10. Sea S un conjunto y sea V un espacio vectorial. Sean F y G dos aplicaciones de S en V. Podemos definir su **suma** de la misma manera en que definimos la suma de funciones, a saber, la suma $F+G$ es la aplicación cuyo valor en un elemento t de S es $F(t)+G(t)$. También definimos el **producto** de F por un número c como la aplicación cuyo valor en un elemento t de S es $cF(t)$. Es fácil verificar que se satisfacen las condiciones **EV 1** hasta la **EV 8**.

Ejercicios IV, §1

1. En el Ejemplo 4, muestre Df como función de x cuando f es la función:
   (a) $f(x) = \operatorname{sen} x$ (b) $f(x) = e^x$ (c) $f(x) = \log x$

2. Sean $P = (0,1)$ y R la rotación determinada por $\pi/4$. Dé las coordenadas de la imagen de P bajo R, es decir, dé $R(P)$.

3. En el Ejemplo 5, dé $L(X)$ cuando X es el vector:
   (a) $(1,2,-3)$ (b) $(-1,5,0)$ (c) $(2,1,1)$

4. Sea $F: \mathbf{R} \to \mathbf{R}^2$ la aplicación tal que $F(t) = (e^t, t)$. ¿A qué son iguales $F(1)$, $F(0)$ y $F(-1)$?

5. Sea $G: \mathbf{R} \to \mathbf{R}^2$ la aplicación tal que $G(t) = (t, 2t)$. Sea F como en el Ejercicio 4. ¿A qué son iguales $(F+G)(1)$, $(F+G)(2)$ y $(F+G)(0)$?

6. Sea F como el Ejercicio 4. ¿A qué son iguales $(2F)(0)$, $(\pi F)(1)$?

7. Sean $A = (1,1,-1,3)$ y $F: \mathbf{R}^4 \to \mathbf{R}$ la aplicación tal que, para cualquier vector $X = (x_1, x_2, x_3, x_4)$, tenemos que $F(X) = X \cdot A + 2$. ¿Cuál es el valor de $F(X)$ cuando (a) $X = (1,1,0,-1)$ y (b) $X = (2,3,-1,1)$?

En los ejercicios 8 a 12, consulte el Ejemplo 6. En cada caso, para probar que la imagen es igual a cierto conjunto S, el lector debe probar que la imagen está contenida en S y también que todo elemento de S está en la imagen.

8. Sea $F: \mathbf{R}^2 \to \mathbf{R}^2$ la aplicación definida por $F(x,y) = (2x, 3y)$. Describa la imagen de los puntos que pertenecen al círculo $x^2 + y^2 = 1$.

9. Sea $F: \mathbf{R}^2 \to \mathbf{R}^2$ la aplicación definida por $F(x,y) = (xy, y)$. Describa la imagen bajo F de la recta $x = 2$.

10. Sea F la aplicación definida por $F(x,y) = (e^x \cos y, e^x \operatorname{sen} y)$. Describa la imagen bajo F de la recta $x = 1$. Describa en forma más general la imagen bajo F de una recta $x = c$, donde c es una constante.

11. Sea F la aplicación definida por $F(t,u) = (\cos t, \operatorname{sen} t, u)$. Describa geométricamente la imagen del plano (t, u) bajo F.

12. Sea F la aplicación definida por $F(x,y) = (x/3, y/4)$. ¿Cuál es la imagen bajo F de la elipse
$$\frac{x^2}{9} + \frac{y^2}{16} = 1?$$

IV, §2. Aplicaciones lineales

Sean V y W dos espacios vectoriales. Una **aplicación lineal**
$$L: V \to W$$
es una aplicación que satisface las siguientes dos propiedades. Primero, para cualesquiera elementos u y v de V, y cualquier escalar c, tenemos

AL 1. $$L(u+v) = L(u) + L(v).$$

AL 2. $$L(cu) = cL(u).$$

Ejemplo 1. La aplicación lineal más importante de este curso queda descrita de la manera siguiente. Sea A una matriz dada de $m \times n$. Defina
$$L_A: \mathbf{R}^n \to \mathbf{R}^m$$
mediante la fórmula
$$L_A(X) = AX.$$
Entonces L_A es lineal. Ciertamente, esto no es otra cosa sino una manera resumida de expresar las propiedades
$$A(X+Y) = AX + AY \qquad \text{y} \qquad A(cX) = cAX$$
para cualesquiera X y Y *verticales* en $\mathbf{R}^n$ y cualquier número c.

Ejemplo 2. El **producto interior** es, esencialmente, un caso especial del primer ejemplo. Sea $A = (a_1, \ldots, a_n)$ un vector fijo, y defina
$$L_A(X) = A \cdot X.$$
Entonces L_A es una aplicación lineal de $\mathbf{R}^n$ en $\mathbf{R}$, debido a que
$$A \cdot (X+Y) = A \cdot X + A \cdot Y \qquad \text{y} \qquad A \cdot (cX) = c(A \cdot X).$$
Observe que el producto interior también se puede considerar como una multiplicación de matrices si se considera A como un vector renglón y X como un vector *columna*.

Ejemplo 3. Sea V cualquier espacio vectorial. La aplicación que asocia a cualquier elemento u de V este mismo elemento es, obviamente, una aplicación lineal que se llama aplicación **identidad**. Denotémosla con I. Así, $I(u) = u$.

Ejemplo 4. Sean V y W cualesquiera espacios vectoriales. La aplicación que asocia el elemento O de W a todo elemento u de V se conoce como aplicación nula y, obviamente, es lineal.

Ejemplo 5. Sea V el conjunto de funciones que tienen derivadas de todos los órdenes. Entonces la derivada $D: V \to V$ es una aplicación lineal. Esto no es sino una manera breve de resumir las propiedades estándar de la derivada, a saber,

$$D(f+g) = Df + Dg,$$
$$D(cf) = cD(f).$$

Ejemplo 6. Sean $V = \mathbf{R}^3$ y $V' = \mathbf{R}^2$ el espacio vectorial de los vectores en el espacio de 3 dimensiones y de 2 dimensiones, respectivamente. Podemos definir una aplicación

$$F: \mathbf{R}^3 \to \mathbf{R}^2$$

mediante la proyección, a saber, $F(x, y, z) = (x, y)$. Dejamos que el lector compruebe que se satisfacen las condiciones **AL 1** y **AL 2**.

En forma más general, suponga que $n = r + s$ se expresa como una suma de dos enteros positivos. Podemos separar las coordenadas $(x_1, \ldots, x_n)$ en dos partes $(x_1, \ldots, x_r, x_{r+1}, \ldots, x_{r+s})$, a saber, las primeras r coordenadas y las últimas s coordenadas. Sea

$$F: \mathbf{R}^n \to \mathbf{R}^r$$

la aplicación tal que $F(x_1, \ldots, x_n) = (x_1, \ldots, x_r)$. Entonces el lector puede verificar fácilmente que F es lineal. Se conoce a F como la **proyección sobre las primeras r coordenadas**. En forma análoga, tendríamos una **proyección sobre las últimas s coordenadas** por medio de la aplicación lineal L tal que

$$L(x_1, \ldots, x_n) = (x_{r+1}, \ldots, x_n).$$

Ejemplo 7. En el cálculo de varias variables se define el gradiente de una función f de la siguiente manera:

$$\operatorname{grad} f(X) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right).$$

Entonces, para dos funciones f y g, tenemos

$$\operatorname{grad}(f+g) = \operatorname{grad} f + \operatorname{grad} g$$

y para cualquier número c,

$$\operatorname{grad}(cf) = c \cdot \operatorname{grad} f.$$

Por tanto, grad es una aplicación lineal.

Sea $L: V \to W$ una aplicación lineal. Sean u, v y w elementos de V. Entonces

$$L(u+v+w) = L(u) + L(v) + L(w).$$

Esto se puede ver por pasos, usando la definición de aplicaciones lineales. Así,

$$L(u+v+w) = L(u+v) + L(w) = L(u) + L(v) + L(w).$$

De la misma manera, dada una suma de más de tres elementos, se satisface una propiedad análoga. Por ejemplo, sean $u_1, \dots, u_n$ elementos de V. Entonces

$$L(u_1 + \cdots + u_n) = L(u_1) + \cdots + L(u_n).$$

La suma de la derecha se puede efectuar en cualquier orden. Se puede dar fácilmente una prueba formal mediante inducción, la cual omitimos.

Si $a_1, \dots, a_n$ son números, entonces

$$L(a_1u_1 + \cdots + a_nu_n) = a_1L(u_1) + \cdots + a_nL(u_n).$$

Mostremos este hecho para tres elementos:

$$\begin{aligned} L(a_1u + a_2v + a_3w) &= L(a_1u) + L(a_2v) + L(a_3w) \\ &= a_1L(u) + a_2L(v) + a_3L(w). \end{aligned}$$

Empleando la notación de suma escribiríamos

$$L\left(\sum_{i=1}^{n} a_iu_i\right) = \sum_{i=1}^{n} a_iL(u_i).$$

En la práctica, las siguientes propiedades se satisfarán de manera obvia, aunque resulta que se pueden probar a partir de los axiomas de aplicaciones lineales y espacios vectoriales.

AL 3. *Sea $L: V \to W$ una aplicación lineal. Entonces $L(O) = O$.*

Demostración. Tenemos que

$$L(O) = L(O+O) = L(O) + L(O).$$

Al restar $L(O)$ en ambos miembros de la igualdad se obtiene $O = L(O)$, tal como se deseaba.

AL 4. *Sea $L : V \mapsto W$ una aplicación lineal. Entonces $L(-v) = -L(v)$.*

Demostración. Tenemos que

$$O = L(O) = L(v-v) = L(v) + L(-v).$$

Al sumar $-L(v)$ a ambos lados de la igualdad obtenemos el aserto deseado.

Podemos observar que los valores de una aplicación lineal quedan determinados si conocemos los valores de los elementos de una base.

Ejemplo 8. Sea $L: \mathbf{R}^2 \to \mathbf{R}^2$ una aplicación lineal. Suponga que

$$L(1,1) = (1,4) \qquad \text{y} \qquad L(2,-1) = (-2,3).$$

Encuentre $L(3,-1)$.

Para hacer esto, escribamos $(3,-1)$ como combinación lineal de $(1,1)$ y $(2,-1)$, de ahí que hemos tenido que resolver

$$(3,-1) = x(1,1) + y(2,-1).$$

Esto equivale a resolver el sistema

$$\begin{aligned} x + 2y &= 3, \\ x - y &= -1. \end{aligned}$$

La solución es $x = \frac{1}{3}$, $y = \frac{4}{3}$. En consecuencia,

$$L(3,-1) = xL(1,1) + yL(2,-1) = \frac{1}{3}(1,4) + \frac{4}{3}(-2,3) = \left(\frac{-7}{3}, \frac{16}{3}\right).$$

Ejemplo 9. Sea V un espacio vectorial y sea $L: V \to \mathbf{R}$ una aplicación lineal. Afirmamos que el conjunto S de todos los elementos v de V, tales que $L(v) < 0$, es convexo.

Demostración. Sean $L(v) < 0$ y $L(w) < 0$. Sea $0 < t < 1$. Entonces

$$L\big(tv + (1-t)w\big) = tL(v) + (1-t)L(w).$$

Entonces $tL(v) < 0$ y $(1,-t)L(w) < 0$, de manera que $tL(v) + (1,-t)L(w) < 0$, por lo que $tv + (1-t)w$ está en S. Si $t = 0$ o $t = 1$, entonces $tv + (1-t)w$ es igual a v o a w y éste también está en S. Esto prueba nuestro aserto.

Consulte el Ejercicio 14 para ver una generalización de este ejemplo.

Las coordenadas de una aplicación lineal

Primero, sea

$$F: V \to \mathbf{R}^n$$

una aplicación cualquiera, de manera que cada valor $F(v)$ es un elemento de $\mathbf{R}^n$ y por tanto, tiene coordenadas. Así, podemos escribir

$$F(v) = \big(F_1(v), \ldots, F_n(v)\big), \qquad \text{o} \qquad F = (F_1, \ldots, F_n).$$

Cada F_i es una función de V en $\mathbf{R}$, que escribimos

$$F_i: V \to \mathbf{R}.$$

Ejemplo 10. Sea $F: \mathbf{R}^2 \to \mathbf{R}^3$ la aplicación

$$F(x,y) = (2x - y, 3x + 4y, x - 5y).$$

Entonces

$$F_1(x,y) = 2x - y, \qquad F_2(x,y) = 3x + 4y, \qquad F_3(x,y) = x - 5y.$$

Observe que cada función coordenada se puede expresar en términos de un producto interior. Por ejemplo, sea

$$A_1 = (2,-1), \qquad A_2 = (3,4), \qquad A_3 = (1,-5).$$

Entonces

$$F_i(x,y) = A_i \cdot (x,y) \qquad \text{para} \qquad i = 1,2,3.$$

Cada función

$$X \mapsto A_i \cdot X$$

es lineal. En forma bastante más general:

Proposición 2.1. *Sea $F: V \to \mathbf{R}^n$ una aplicación de un espacio vectorial V en $\mathbf{R}^n$. Entonces F es lineal si, y sólo si, cada función coordenada $F_i: V \to \mathbf{R}$ es lineal, para $i = 1, \ldots, n$.*

Demostración. Para v y $w \in V$, tenemos

$$F(v+w) = \big(F_1(v+w), \ldots, F_n(v+w)\big),$$
$$F(v) = \big(F_1(v), \ldots, F_n(v)\big),$$
$$F(w) = \big(F_1(w), \ldots, F_n(w)\big),$$

por tanto, $F(v+w) = F(v) + F(w)$ si, y sólo si, $F_i(v+w) = F_i(v) + F_i(w)$ para todo $i = 1, \ldots, n$ por la definición de adición de n-tuplas. El mismo argumento demuestra que, si $c \in \mathbf{R}$, entonces $F(cv) = cF(v)$ si, sólo si,

$$F_i(cv) = cF_i(v) \qquad \text{para todo} \qquad i = 1, \ldots, n.$$

Esto prueba la proposición.

Ejemplo 10 (continúa). La aplicación del Ejemplo 10 es lineal debido a que todas las funciones coordenadas son lineales. En efecto, si el lector escribe el vector (x,y) en forma vertical, se dará cuenta de que la aplicación F es, de hecho, igual a L_A para alguna matriz A. ¿Cuál es esta matriz A?

El espacio vectorial de las aplicaciones lineales

Sean V y W dos espacios vectoriales. Consideremos el conjunto de todas las aplicaciones lineales de V en W y denotemos este conjunto con $\mathcal{L}(V,W)$, o simplemente $\mathcal{L}$ si la referencia a V y W es clara. Definiremos la adición de aplicaciones lineales y su multiplicación por números de tal manera que $\mathcal{L}$ se convierta en un espacio vectorial.

Sean $L: V \to W$ y $F: V \to W$ dos aplicaciones lineales. Definimos su **suma** $L + F$ como la aplicación cuyo valor en un elemento u de V es $L(u) + F(u)$. Por tanto, podemos escribir

$$(L+F)(u) = L(u) + F(u).$$

Entonces la aplicación $L + F$ es lineal. Ciertamente, es fácil verificar que se cumplen las dos condiciones que definen una aplicación lineal. Para cualesquiera elementos u y v de V, tenemos

$$\begin{aligned}(L+F)(u+v) &= L(u+v)+F(u+v)\\ &= L(u)+L(v)+F(u)+F(v)\\ &= L(u)+F(u)+L(v)+F(v)\\ &= (L+F)(u)+(L+F)(v).\end{aligned}$$

Además, si c es un número, entonces

$$\begin{aligned}(L+F)(cu) &= L(cu)+F(cu)\\ &= cL(u)+cF(u)\\ &= c[L(u)+F(u)]\\ &= c[(L+F)(u)].\end{aligned}$$

En consecuencia, $L + F$ es una aplicación lineal.

Si a es un número y $L: V \to W$ es una aplicación lineal, entonces definimos una aplicación aL de V en W dando su valor en un elemento u de V, a saber, $(aL)(u) = aL(u)$. Entonces se verifica con facilidad que aL es una aplicación lineal; dejamos esto como ejercicio.

Acabamos de definir operaciones de adición y multiplicación por números en nuestro conjunto $\mathcal{L}$. Además, si $L: V \to W$ es una aplicación lineal, i.e., un elemento de $\mathcal{L}$, entonces podemos definir $-L$ como $(-1)L$, es decir, el producto del número -1 por L. Por último, tenemos la **aplicación nula**, que a cada elemento de V le asocia el elemento O de V . Entonces $\mathcal{L}$ es un espacio vectorial. En otras palabras, el conjunto de aplicaciones lineales de V en W es él mismo un espacio vectorial. La verificación de que se satisfacen las reglas **EV 1** a **EV 8** para un espacio vectorial es sencilla y se deja para el lector.

Ejemplo 11. Sea $V = W$ el espacio vectorial de funciones que tienen derivadas de todos los órdenes. Sea D la derivada y sea I la identidad. Si f está en V, entonces

$$(D+I)f = Df + f.$$

Así, cuando $f(x) = e^x$, entonces $(D+I)f$ es la función cuyo valor en x es $e^x + e^x = 2e^x$.

Si $f(x) = \operatorname{sen} x$, entonces $(D+3I)f$ es la función tal que

$$((D+3I)f)(x) = (Df)(x) + 3If(x) = \cos x + 3\operatorname{sen} x.$$

Se observa que $3 \cdot I$ es una aplicación lineal, cuyo valor en f es $3f$. Así, $(D + 3\cdot I)f = Df + 3f$. En cualquier número x, el valor de $(D + 3\cdot I)f$ es $Df(x) + 3f(x)$. También podemos escribir $(D+3I)f = Df + 3f$.

Ejercicios IV, §2

1. Determine cuáles de las siguientes aplicaciones F son lineales.
   (a) $F: \mathbf{R}^3 \to \mathbf{R}^2$ determinada por $F(x, y, z) = (x, z)$.
   (b) $F: \mathbf{R}^4 \to \mathbf{R}^4$ determinada por $F(X) = -X$.
   (c) $F: \mathbf{R}^3 \to \mathbf{R}^3$ determinada por $F(X) = X + (0, -1, 0)$.
   (d) $F: \mathbf{R}^2 \to \mathbf{R}^2$ determinada por $F(x, y) = (2x + y, y)$.
   (e) $F: \mathbf{R}^2 \to \mathbf{R}^2$ determinada por $F(x, y) = (2x, y - x)$.
   (f) $F: \mathbf{R}^2 \to \mathbf{R}^2$ determinada por $F(x, y) = (y, x)$.
   (g) $F: \mathbf{R}^2 \to \mathbf{R}$ determinada por $F(x, y) = xy$.

2. ¿Cuáles de las aplicaciones que aparecen en los Ejercicios 4, 7, 8, 9 de la sección §1, son lineales?

3. Sean V y W dos espacios vectoriales y sea $F: V \to W$ una aplicación lineal. Sea U el subconjunto de V que consta de todos los elementos v tales que $F(v) = O$. Pruebe que U es un subespacio de V.

4. Sea $L: V \to W$ una aplicación lineal. Pruebe que la imagen de L es un subespacio de W. [Esto se efectuará en la siguiente sección, pero inténtelo ahora el lector, para que adquiera práctica.]

5. Sean A y B dos matrices de $m \times n$. Suponga que

$$AX = BX$$

para todas las n-tuplas X. Muestre que $A = B$. Esto también se puede enunciar en la siguiente forma: si $L_A = L_B$, entonces $A = B$.

6. Sea $T_u: V \to V$ la traslación determinada por un vector u. ¿Para qué vectores u es una aplicación lineal T_u?

7. Sea $L: V \to W$ una aplicación lineal.
   (a) Si S es una recta en V, muestre que la imagen $L(S)$ es, o bien una recta en W, o bien un punto.
   (b) Si S es un segmento de recta en V entre los puntos P y Q, demuestre que la imagen $L(S)$ es, o bien un punto, o bien un segmento de recta en W. ¿Entre cuáles puntos de W?
   (c) Sean v_1 y v_2 elementos de V linealmente independientes. Suponga que $L(v_1)$ y $L(v_2)$ son linealmente independientes en W. Sea P un elemento de V y sea S el paralelogramo

$$P + t_1 v_1 + t_2 v_2 \qquad \text{donde} \qquad 0 \leq t_i \leq 1 \qquad \text{para} \qquad i = 1, 2.$$

   Muestre que la imagen $L(s)$ es un paralelogramo en W.
   (d) Sean v y w elementos linealmente independientes de un espacio vectorial V. Sea $F: V \to W$ una aplicación lineal. Suponga que $F(v)$ y $F(w)$ son linealmente dependientes. Demuestre que la imagen bajo F del paralelogramo generado por v y w es un punto, o bien un segmento de recta.

8. Sean $E_1 = (1,0)$ y $E_2 = (0,1)$ como de costumbre. Sea F una aplicación lineal de $\mathbf{R}^2$ en sí mismo, tal que

$$F(E_1) = (1,1) \qquad \text{y} \qquad F(E_2) = (-1,2).$$

Sea S el cuadrado con vértices en $(0,0)$, $(1,0)$, $(1,1)$ y $(0,1)$. Muestre que la imagen de este cuadrado bajo F es un paralelogramo.

9. Sean A y B dos vectores no nulos en el plano tales que no hay ninguna constante $c \neq 0$ tal que $B = cA$. Sea L una aplicación lineal del plano en sí mismo tal que $L(E_1) = A$ y $L(E_2) = B$. Describa la imagen bajo L del rectángulo cuyos vértices se encuentran en $(0,1)$, $(3,0)$, $(0,0)$ y $(3,1)$.

10. Sea $L\colon \mathbf{R}^2 \to \mathbf{R}^2$ una aplicación lineal, que tiene el siguiente efecto sobre los vectores indicados:
(a) $L(3,1) = (1,2)$ y $L(-1,0) = (1,1)$
(b) $L(4,1) = (1,1)$ y $L(1,1) = (3,-2)$
(c) $L(1,1) = (2,1)$ y $L(-1,1) = (6,3)$.
En cada uno de los casos calcule $L(1,0)$.

11. Sea L como en (a), (b) y (c) del ejercicio 10. Encuentre $L(0,1)$.

12. Sean V y W dos espacios vectoriales y $F\colon V \to W$ una aplicación lineal. Sean $w_1, \dots, w_n$ elementos de W, los que son linealmente independientes, y sean $v_1, \dots, v_n$ elementos de V tales que $F(v_i) = w_i$ para $i = 1, \dots, n$. Demuestre que $v_1, \dots, v_n$ son linealmente independientes.

13. (a) Sea V un espacio vectorial y $F\colon V \to \mathbf{R}$ una aplicación lineal. Sea W el subconjunto de V que consta de todos los elementos v tales que $F(v) = 0$. Suponga que $W \neq V$ y sea v_0 un elemento de V que no está en W. Demuestre que todo elemento de V se puede escribir como una suma $w + cv_0$, con algún w en W y algún número c.
(b) Muestre que W es un subespacio de V. Sea $\{v_1, \dots, v_n\}$ una base de W. Muestre que $\{v_0, v_1, \dots, v_n\}$ es una base de V.

Conjuntos convexos

14. Demuestre que la imagen de un conjunto convexo bajo una aplicación lineal es convexa.

15. Sea $L\colon V \to W$ una aplicación lineal. Sea T un conjunto convexo en W y sea S el conjunto de elementos $v \in V$ tales que $L(v) \in T$. Demuestre que S es convexo.

Observación. ¿Por qué estos ejercicios dan una prueba más general que la que el lector debería haber desarrollado con anterioridad? Por ejemplo: sea $A \in \mathbf{R}^n$ y sea c un número. Entonces el conjunto de todos los $X \in \mathbf{R}^n$ tales que $X \cdot A \geq c$ es convexo. Además, si S es un conjunto convexo y c es un número, entonces cS es convexo. ¿Cómo se interpretan estos enunciados como casos especiales de los Ejercicios 14 y 15?

16. Sea S un conjunto convexo en V y sea $u \in V$. Sea $T_u\colon V \to V$ la traslación determinada por u. Muestre que la imagen de $T_u(S)$ es convexa.

Vectores propios y valores propios. Sea V un espacio vectorial y sea $L: V \to V$ una aplicación lineal. Un **vector propio** v para L es un elemento de V tal que existe un escalar c con la propiedad siguiente:

$$L(v) = cv$$

El escalar c se conoce como **valor propio** de v con respecto a L. Si $v \neq 0$, entonces c está determinado en forma única. Cuando V es un espacio vectorial cuyos elementos son funciones, entonces a un vector propio también se le conoce como **función propia**.

17. (a) Sea V el espacio de las funciones diferenciables en $\mathbf{R}$. Sea $f(t) = e^{ct}$, donde c es algún número. Sea L la derivada d/dt. Demuestre que f es una función propia para L. ¿Cuál es el valor propio?
    (b) Sea L segunda derivada, esto es,

$$L(f) = \frac{d^2 f}{dt^2}$$

para cualquier función f. Demuestre que las funciones $\operatorname{sen} t$ y $\cos t$ son funciones propias de L. ¿Cuáles son los valores propios?

18. Sea $L: V \to V$ una aplicación lineal, y sea W el subconjunto de elementos de V formado por todos los vectores propios de L con un valor propio c dado. Muestre que W es un subespacio.

19. Sea $L: V \to V$ una aplicación lineal. Sean $v_1, \ldots, v_n$ vectores propios para L, con valores propios $c_1, \ldots, c_n$, respectivamente. Suponga que $c_1, \ldots, c_n$ son distintos entre sí. Pruebe que $v_1, \ldots, v_n$ son linealmente independientes. [*Sugerencia*: Use inducción.]

IV, §3. El núcleo y la imagen de una aplicación lineal

Sea $F: V \to W$ una aplicación lineal. La **imagen** de F es el conjunto de elementos w de W para los cuales existe un elemento v de V tal que $F(v) = w$.

La imagen de F es un subespacio de W.

Demostración. Observe primero que $F(O) = O$ y, en consecuencia, que O está en la imagen. Luego suponga que w_1 y w_2 están en la imagen. Entonces existen elementos v_1 y v_2 de V tales que $F(v_1) = w_1$ y $F(v_2) = w_2$. Por tanto,

$$F(v_1 + v_2) = F(v_1) + F(v_2) = w_1 + w_2,$$

con lo que se prueba que $w_1 + w_2$ está en la imagen. Si c es un número, entonces

$$F(cv_1) = cF(v_1) = cw_1.$$

Por tanto, cw_1 está en la imagen. Esto prueba que la imagen es un subespacio de W.

Sean V y W espacios vectoriales y sea $F: V \to W$ una aplicación lineal. El conjunto de elementos $v \in V$ tales que $F(v) = 0$ se conoce como **núcleo** de F.

El núcleo de F es un subespacio de V.

Demostración. Como $F(O) = O$, vemos que O está en el núcleo. Supongamos que v y w están en el núcleo. Entonces $F(v + w) = F(v) + F(w) = O + O = O$, de manera que $v + w$ está en el núcleo. Si c es un número, entonces $F(cv) = cF(v) = O$, por lo que cv también está en el núcleo. En consecuencia, el núcleo es un subespacio.

Ejemplo 1. Sea $L: \mathbf{R}^3 \to \mathbf{R}$ la aplicación tal que

$$L(x, y, z) = 3x - 2y + z.$$

Así, si $A = (3, -2, 1)$, podemos escribir

$$L(X) = X \cdot A = A \cdot X.$$

Entonces el núcleo de L es el conjunto de soluciones de la ecuación

$$3x - 2y + z = 0.$$

Por supuesto, esto se generaliza al espacio de dimensión n. Si A es un vector arbitrario de $\mathbf{R}^n$, podemos definir la aplicación lineal

$$L_A: \mathbf{R}^n \to \mathbf{R}$$

tal que $L_A(X) = A \cdot X$. Su núcleo se puede interpretar como el conjunto de todos los X que son perpendiculares a A.

Ejemplo 2. Sea $P: \mathbf{R}^3 \to \mathbf{R}^2$ la proyección, tal que

$$P(x, y, z) = (x, y).$$

Entonces P es una aplicación lineal cuyo núcleo consta de todos los vectores de $\mathbf{R}^3$ cuyas primeras dos coordenadas son iguales a 0, i.e., todos los vectores

$$(0, 0, z),$$

donde z es una componente arbitraria.

Ejemplo 3. Sea A una matriz de $m \times n$ y sea

$$L_A: \mathbf{R}^n \to \mathbf{R}^m$$

la aplicación lineal tal que $L_A(X) = AX$. Entonces el núcleo de L_A es, precisamente, el subespacio de soluciones X de las ecuaciones lineales

$$AX = O.$$

Ejemplo 4. Ecuaciones diferenciales. Sea D la derivada. Si la variable real se denota con x, entonces también podemos escribir $D = d/dx$. La derivada se puede iterar, de manera que la segunda derivada se denota con D^2 (o $(d/dx)^2$). Cuando se aplica a una función, escribimos $D^2 f$, de manera que

$$(D^2 f)(x) = \frac{d^2 f}{dx^2}.$$

Se procede en forma análoga para D^3, $D^4, \ldots, D^n$ para la derivada de orden n.

Ahora sea V el espacio vectorial de funciones que admiten derivadas de todos los órdenes. Sean $a_1, \ldots, a_m$ números, y sea g un elemento de V, esto es, una función infinitamente diferenciable. Considere el problema de hallar una solución f de la ecuación diferencial

$$a_m \frac{d^m f}{dx^m} + a_{m-1} \frac{d^{m-1} f}{dx^{m-1}} + \cdots + a_1 f = g.$$

Podemos reescribir esta ecuación sin la variable x, en la forma

$$a_m D^m f + a_{m-1} D^{m-1} f + \cdots + a_1 f = g.$$

Cada derivada D^k es una aplicación lineal de V en sí mismo. Sea

$$L = a_m D^m + a_{m-1} D^{m-1} + \cdots + a_1 I.$$

Entonces L es una suma de aplicaciones lineales y es ella misma una aplicación lineal. Así, la ecuación diferencial se puede escribir en la forma siguiente:

$$L(f) = g.$$

Ésta se encuentra ahora en una notación similar a la que se usó para resolver ecuaciones lineales. Además, esta ecuación está en forma "no homogénea". La ecuación homogénea asociada es la ecuación

$$L(f) = 0,$$

donde el miembro de la derecha de la igualdad es la función nula. Sea W el núcleo de L. Entonces W es el conjunto (espacio) de soluciones de la ecuación homogénea

$$a_m D^m f + \cdots + a_1 f = 0.$$

Si existe una solución f_0 para la ecuación no homogénea $L(f) = g$, entonces todas las soluciones se obtienen mediante la traslación

$$f_0 + W = \text{conjunto de todas las funciones } f_0 + f, \text{ donde } f \text{ está en } W.$$

Vea el Ejercicio 5.

En varios ejercicios anteriores buscamos la imagen de las rectas, de los planos y de los paralelogramos bajo una aplicación lineal. Por ejemplo, si consideramos el plano generado por dos vectores linealmente independientes v_1 y v_2 de V, y

$$L: V \to W$$

es una aplicación lineal, entonces la imagen de ese plano será un plano, siempre que $L(v_1)$ y $L(v_2)$ también sean linealmente independientes. Podemos dar un criterio para esto en términos del núcleo y el criterio es válido en forma bastante general de la manera siguiente.

Teorema 3.1. *Sea $F: V \to W$ una aplicación lineal cuyo núcleo es $\{O\}$. Si $v_1, \ldots, v_n$ son elementos linealmente independientes de V, entonces $F(v_1), \ldots, F(v_n)$ son elementos de W linealmente independientes.*

Demostración. Sean $x_1, \ldots, x_n$ números tales que

$$x_1 F(v_1) + \cdots + x_n F(v_n) = O.$$

Debido a la linealidad, tenemos que

$$F(x_1 v_1 + \cdots + x_n v_n) = O.$$

En consecuencia, $x_1 v_1 + \cdots + x_n v_n = O$. Puesto que $v_1, \ldots, v_n$ son linealmente independientes, se infiere que $x_i = 0$ para $i = 1, \ldots, n$. Esto prueba nuestro teorema.

A menudo abreviamos núcleo e imagen escribiendo Ker e Im, respectivamente. El siguiente teorema relaciona las dimensiones del núcleo y de la imagen de una aplicación lineal, con la dimensión del espacio en el cual está definida la aplicación.

Teorema 3.2. *Sea V un espacio vectorial. Sea $L: V \to W$ una aplicación lineal de V en otro espacio W. Sea n la dimensión de V, q la dimensión del núcleo de L y s la dimensión de la imagen de L. Entonces $n = q + s$. En otras palabras,*

$$\dim V = \dim \operatorname{Ker} L + \dim \operatorname{Im} L.$$

Demostración. Si la imagen de L consta sólo del O, entonces nuestro aserto es trivial. Por consiguiente, podemos suponer que $s > 0$. Sea $\{w_1, \ldots, w_s\}$ una base de la imagen de L. Sean $v_1, \ldots, v_s$ elementos de V tales que $L(v_i) = w_i$ para $i = 1, \ldots, s$. Si el núcleo no es $\{O\}$, sea $\{u_1, \ldots, u_q\}$ una base del núcleo. Si el núcleo es $\{O\}$, deberá quedar entendido que en lo que sigue será omitida toda referencia a $\{u_1, \ldots, u_q\}$. Afirmamos que

$$\{v_1, \ldots, v_s, u_1, \ldots, u_q\}$$

es una base de V. Esto será suficiente para probar nuestro aserto. Sea v cualquier elemento de V. Entonces existen números $x_1, \ldots, x_s$ tales que

$$L(v) = x_1 w_1 + \cdots + x_s w_s,$$

debido a que $\{w_1, \ldots, w_s\}$ es una base de la imagen de L. Por linealidad,

$$L(v) = L(x_1 v_1 + \cdots + x_s v_s),$$

y, de nuevo por linealidad, al restar el miembro derecho del miembro izquierdo se infiere que

$$L(v - x_1 v_1 - \cdots - x_s v_s) = O.$$

En consecuencia, $v - x_1 v_1 - \cdots - x_s v_s$ está en el núcleo de L y existen números $y_1, \ldots, y_q$ tales que

$$v - x_1 v_1 - \cdots - x_s v_s = y_1 u_1 + \cdots + y_q u_q.$$

Por tanto,

$$v = x_1 v_1 + \cdots + x_s v_s + y_1 u_1 + \cdots + y_q u_q$$

es una combinación lineal de $v_1, \ldots, v_s$, $u_1, \ldots, u_q$, lo cual pruebá que estos $s+q$ elementos de V generan V.

Ahora mostramos que son linealmente independientes y, en consecuencia, que constituyen una base. Suponga que existe una relación lineal:

$$x_1v_1 + \cdots + x_sv_s + y_1u_1 + \cdots + y_qu_q = O.$$

Al aplicar L a esta relación y al usar el hecho de que $L(u_j) = O$ para $j = 1, \ldots, q$, obtenemos

$$x_1L(v_1) + \cdots + x_sL(v_s) = O.$$

Pero $L(v_1), \ldots, L(v_s)$ no son otra cosa más que $w_1, \ldots, w_s$, los cuales se supone que son linealmente independientes. De modo que $x_i = 0$ para $i = 1, \ldots, s$, y por ello,

$$y_1u_1 + \cdots + y_qu_q = 0.$$

Pero $u_1, \ldots, u_q$ constituye una base del núcleo de L y, en particular, son linealmente independientes, por lo que todos los $y_j = 0$ para $j = 1, \ldots, q$. Esto concluye la demostración de nuestro aserto.

Ejemplo 1 (continúa). La aplicación lineal $L: \mathbf{R}^3 \to \mathbf{R}$ del Ejemplo 1 está dada mediante la fórmula

$$L(x, y, z) = 3x - 2y + z.$$

Su núcleo consta de todas las soluciones de la ecuación

$$3x - y + z = 0.$$

Su imagen es un subespacio de $\mathbf{R}$, no es $\{O\}$ y, por tanto, es todo $\mathbf{R}$. Así pues, su imagen tiene dimensión 1 y, en consecuencia, su núcleo tiene dimensión 2.

Ejemplo 2 (continúa). La imagen de la proyección

$$P: \mathbf{R}^3 \to \mathbf{R}^2$$

del Ejemplo 2 es igual a todo $\mathbf{R}^2$ y el núcleo tiene dimensión igual a 1.

Ejercicios IV, §3

Sea $L: V \to W$ una aplicación lineal.

1. (a) Si S es un subespacio de V de una dimensión, demuestre que la imagen $L(S)$ es, o bien un punto, o bien una recta.
   (b) Si S es un subespacio de V de dos dimensiones, demuestre que la imagen $L(S)$ puede ser un plano, una recta o un punto.
2. (a) Si S es una recta arbitraria de V (consúltese el Capítulo III, sección §2), muestre que la imagen de S es, o bien un punto, o bien una recta.
   (b) Si S es un plano arbitrario de V, demuestre que la imagen de S puede ser un plano, una recta o un punto.

3. (a) Sea $F: V \to W$ una aplicación lineal cuyo núcleo es $\{O\}$. Suponga que V y W tienen la misma dimensión n. Demuestre que la imagen de F es todo W.
   (b) Sea $F: V \to W$ una aplicación lineal y suponga que la imagen de F es todo W. Suponga que V y W tienen la misma dimensión n. Muestre que el núcleo de F es $\{O\}$.

4. Sea $L: V \to W$ una aplicación lineal. Suponga que $\dim V > \dim W$. Muestre que el núcleo de L no es O.

5. Sea $L: V \to W$ una aplicación lineal. Sea w un elemento de W. Sea v_0 un elemento de V tal que $L(v_0) = w$. Muestre que cualquier solución de la ecuación $L(X) = w$ es del tipo $v_0 + u$, donde u es un elemento del núcleo de L.

6. Sea V el espacio vectorial de las funciones que tienen derivadas de todos los órdenes, y sea $D: V \to V$ la derivada. ¿Cuál es el núcleo de D?

7. Sea D^2 la segunda derivada (i.e., la iteración de D tomada dos veces). ¿Cuál es el núcleo de D^2? En general, ¿cuál es el núcleo de D^n (la n-ésima derivada)?

8. (a) Sean V y D como en el Ejercicio 6. Sea $L = D - I$, donde I es la aplicación identidad de V. ¿Cuál es el núcleo de L?
   (b) La misma pregunta para $L = D - aI$, donde a es un número.

9. (a) ¿Cuál es la dimensión del subespacio de $\mathbf{R}^n$ que consta de aquellos vectores $A = (a_1, \ldots, a_n)$ tales que $a_1 + \cdots + a_n = 0$?
   (b) ¿Cuál es la dimensión del subespacio del espacio de las matrices (a_{ij}) de $n \times n$ tales que

$$a_{11} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii} = 0?$$

10. Se dice que una matriz A de $n \times n$ es **antisimétrica** si ${}^tA = -A$. Demuestre que cualquier matriz A de $n \times n$ se puede escribir como una suma

$$A = B + C,$$

donde B es simétrica y C es antisimétrica. [*Sugerencia*: Sea $B = (A + {}^tA)/2$.] Demuestre que, si $A = B_1 + C_1$, donde B_1 es simétrica y C_1 es antisimétrica, entonces $B = B_1$ y $C = C_1$.

11. Sea M el espacio de todas las matrices de $n \times n$. Sea

$$P: M \to M$$

la aplicación tal que

$$P(A) = \frac{A + {}^tA}{2}.$$

   (a) Demuestre que P es lineal.
   (b) Demuestre que el núcleo de P consta del espacio de las matrices antisimétricas.
   (c) Demuestre que la imagen de P consta de todas las matrices simétricas. [*Cuidado.* Se tiene que probar dos cosas: para cualquier matriz A, $P(A)$ es simétrica. Recíprocamente, dada una matriz simétrica B, existe una matriz A tal que $B = P(A)$. ¿Cuál es la posibilidad más simple para tal A?]
   (d) El lector debió haber determinado con anterioridad la dimensión del espacio de las matrices simétricas y seguramente encontró que era igual a $n(n+1)/2$. ¿Cuál es entonces la dimensión del espacio de las matrices antisimétricas?
   (e) Exhiba una base para el espacio de las matrices antisimétricas.

12. Sea M el espacio de todas las matrices de $n \times n$. Sea

$$Q: M \to M$$

la aplicación tal que

$$Q(A) = \frac{A - {}^tA}{2}.$$

(a) Demuestre que Q es lineal.
(b) Describa el núcleo de Q y determine su dimensión.
(c) ¿Cuál es la imagen de Q?

13. Se dice que una función (evaluada en los reales, de una variable real) es **par** si $f(-x) = f(x)$. Se dice que es **impar** si $f(-x) = -f(x)$.
(a) Verifique que $\operatorname{sen} x$ es una función impar y que $\cos x$ es una función par.
(b) Sea V el espacio vectorial de todas las funciones. Defina la aplicación

$$P: V \to V$$

como $(Pf)(x) = \big(f(x) + f(-x)\big)/2$. Demuestre que P es una aplicación lineal.
(c) ¿Cuál es el núcleo de P?
(d) ¿Cuál es la imagen de P? Pruebe sus asertos.

14. De nuevo, sea V el espacio vectorial de todas las funciones. Defina la aplicación

$$Q: V \to V$$

como $(Qf)(x) = \big(f(x) - f(-x)\big)/2$.
(a) Demuestre que Q es una aplicación lineal.
(b) ¿Cuál es el núcleo de Q?
(c) ¿Cuál es la imagen de Q? Pruebe su aserto.

Observación. Los Ejercicios 11, 12, 13 y 14 tienen ciertos elementos formales en común. Estas características comunes se discutirán más adelante. Consulte los Ejercicios 4 a 7 del Capítulo V, sección §1.

15. **El espacio producto.** Sean U y W espacios vectoriales. El **producto directo**, $U \times W$ que simplemente llamaremos **producto**, es el conjunto de todas las parejas (u, w) con $u \in U$ y $w \in W$. No deberá confundirse con el producto de números, con el producto escalar ni con el producto vectorial, que algunas veces se usa en física para denotar un tipo diferente de operación. Es un hecho histórico desafortunado que la palabra producto se use en dos diferentes contextos, y el lector deberá acostumbrarse a ello. Por ejemplo, podemos considerar $\mathbf{R}^4$ como un producto,

$$\mathbf{R}^4 = \mathbf{R}^3 \times \mathbf{R}^1 = \mathbf{R}^3 \times \mathbf{R}$$

tomando una tetrada (x_1, x_2, x_3, x_4) como el resultado de colocar juntas a la terna (x_1, x_2, x_3) y al número sencillo x_4. En forma análoga,

$$\mathbf{R}^4 = \mathbf{R}^2 \times \mathbf{R}^2,$$

considerando (x_1, x_2, x_3, x_4) como el resultado de colocar juntos a (x_1, x_2) y (x_3, x_4).

Si (u_1, w_1) y (u_2, w_2) son elementos de $U \times W$, de manera que

$$u_1, u_2 \in U \qquad \text{y} \qquad w_1, w_2 \in W,$$

definimos su suma componente a componente, esto es, definimos

$$(u_1, w_1) + (u_2, w_2) = (u_1 + u_2, w_1 + w_2).$$

Si c es un número, definimos $c(u, w) = (cu, cw)$.

(a) Demuestre que $U \times W$ es un espacio vectorial con estas definiciones. ¿Cuál es el elemento cero?

(b) Demuestre que $\dim(U \times W) = \dim U + \dim W$. En efecto, sea $\{u_i\}$ ($i = 1, \ldots, n$) una base de U y $\{w_j\}$ ($j = 1, \ldots, m$) una base de W. Demuestre que los elementos $\{(u_i, 0)\}$ y $\{(0, w_j)\}$ forman una base de $U \times W$.

(c) Sea U un subespacio de un espacio vectorial V. Demuestre que el subconjunto de $V \times V$ que consta de todos los elementos (u, u) con $u \in U$, es un subespacio de $V \times V$.

(d) Sea $\{u_i\}$ una base de U. Demuestre que el conjunto de elementos (u_i, u_i) es una base del subespacio que aparece en (c). En consecuencia, la dimensión de este subespacio es la misma que la dimensión de U.

16. (Debe hacerse después de realizar el Ejercicio 15.) Sean U y W subespacios de un espacio vectorial V. Demuestre mediante el método indicado que

$$\boxed{\dim U + \dim W = \dim(U + W) + \dim(U \cap W).}$$

(a) Demuestre que la aplicación

$$L: U \times W \to V$$

dada por

$$L(u, w) = u - w$$

es una aplicación lineal.

(b) Demuestre que la imagen de L es $U + W$.

(c) Demuestre que el núcleo de L es el subespacio de $U \times W$ que consta de todos los elementos (u, u) donde u está en $U \cap W$. ¿Cuál es una base para este subespacio? ¿Cuál es su dimensión?

(d) Aplique la fórmula de la dimensión que aparece en el texto para concluir la demostración.

IV, §4. El rango y las ecuaciones lineales de nuevo

Sea A una matriz de $m \times n$,

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

Sea $L_A: \mathbf{R}^n \to \mathbf{R}^m$ la aplicación lineal que se definió antes, a saber,

$$L_A(X) = AX.$$

Como ya hemos mencionado, el núcleo de L_A es el espacio de soluciones del sistema de ecuaciones lineales escrito en forma abreviada como

$$AX = O.$$

Analicemos ahora su imagen.

Sean $E^1, \ldots, E^n$ los vectores unitarios canónicos de $\mathbf{R}^n$, escritos como vectores columna de la siguiente manera:

$$E^1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad E^n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Entonces la multiplicación ordinaria de matrices muestra que

$$AE^j = A^j$$

es la j-ésima columna de A. En consecuencia, para cualquier vector

$$X = x_1E^1 + \cdots + x_nE^n,$$

encontramos que

$$AX = L_A(X) = x_1A^1 + \cdots + x_nA^n.$$

Así, vemos que:

Teorema 4.1. *La imagen de L_A es el subespacio generado por las columnas de A.*

En el Capítulo III dimos un nombre a la dimensión de ese espacio, a saber, **rango por columnas**, y vimos que es igual al rango por renglones y que se conoce simplemente como **rango** de A. Ahora podemos interpretar este rango también de la siguiente manera:

El rango de A es la dimensión de la imagen de L_A.

Teorema 4.2. *Sea r el rango de A. Entonces la dimensión del espacio de soluciones de $AX = O$ es igual a $n - r$.*

Demostración. Por el Teorema 3.2 tenemos que

$$\dim \operatorname{Im} L_A + \dim \operatorname{Ker} L_A = n.$$

Pero $\dim \operatorname{Im} L_A = r$ y $\operatorname{Ker} L_A$ es el espacio de soluciones de las ecuaciones lineales homogéneas, de manera que nuestro aserto ahora es claro.

Ejemplo 1. Encuentre la dimensión del espacio de soluciones del sistema de ecuaciones siguiente:

$$\begin{aligned} 2x + y + z + 2w &= 0 \\ x + y - 2z - w &= 0. \end{aligned}$$

En este caso, la matriz A es

$$\begin{pmatrix} 2 & -1 & 1 & 2 \\ 1 & 1 & -2 & -1 \end{pmatrix}.$$

Tiene rango 2 debido a que los dos vectores

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

son linealmente independientes, como se puede ver con facilidad. [Use operaciones por renglón y por columna, o bien, hágalo mediante ecuaciones lineales.] En consecuencia, la dimensión del espacio de soluciones es $4-2=2$.

Recordemos que el sistema de ecuaciones lineales también se podría escribir en la siguiente forma:

$$X \cdot A_i = 0 \qquad \text{para} \qquad i = 1, \ldots, m,$$

donde A_i son los renglones de la matriz A. Esto significa que X es perpendicular a cada renglón de A. Entonces X también es perpendicular al espacio de renglones de A, i.e., al espacio generado por los renglones. Ahora resulta conveniente introducir alguna terminología.

Sea U un subespacio de $\mathbf{R}^n$. Sea

> $U^\perp$ = conjunto de todos los elementos X de $\mathbf{R}^n$ tal que $X \cdot Y = O$ para todo Y en U.

Decimos que $U^\perp$ es el **complemento ortogonal** de U. Es el conjunto de los vectores que son perpendiculares a todos los elementos de U, o bien, como también diremos, perpendicular al propio U. Entonces se verifica con facilidad que $U^\perp$ es un subespacio (Ejercicio 8).

Sea U el subespacio generado por los vectores renglón de la matriz $A = (a_{ij})$. Entonces su complemento ortogonal $U^\perp$ es precisamente el conjunto de soluciones de las ecuaciones homogéneas

$$X \cdot A_i = 0 \qquad \text{para todo } i.$$

En otras palabras, tenemos que

(espacio de renglones de $A)^\perp = \operatorname{Ker} L_A$ = espacio de soluciones de $AX = O$.

Teorema 4.3. *Sea U un subespacio de $\mathbf{R}^n$. Entonces*

$$\dim U + \dim U^\perp = n.$$

Demostración. Sea $r = \dim U$. Si $r = 0$, entonces el aserto es obvio. Si $r \neq 0$, entonces U tiene una base y, en particular, está generada por un número finito de vectores $A_1, \ldots, A_r$, que se pueden considerar como los renglones de una matriz. Entonces la fórmula de la dimensión es un caso especial del Teorema 4.2.

En el espacio de tres dimensiones, por ejemplo, el Teorema 4.3 prueba el hecho de que el complemento ortogonal de una recta es un plano y viceversa, tal como se muestra en la figura.

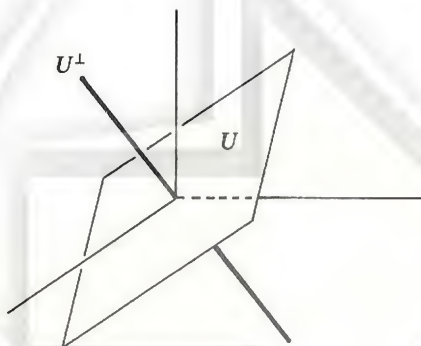


Figura 2

En el espacio de 4 dimensiones, el complemento ortogonal de un subespacio de dimensión 1 tiene dimensión 3. El complemento ortogonal de un espacio de dimensión 2 también tiene dimensión 2.

Estudiemos ahora brevemente las ecuaciones no homogéneas, esto es, un sistema de la forma

$$AX = B,$$

donde B es un vector dado (m-tupla). Dicho sistema quizás no tenga una solución; en otras palabras, las ecuaciones pueden ser lo que se llama "inconsistentes".

Ejemplo 2. Considere el sistema

$$\begin{aligned} 3x - y + z &= 1, \\ 2x + y - z &= 2, \\ x - 2y + 2z &= 5. \end{aligned}$$

Resulta que el tercer renglón de la matriz de coeficientes

$$A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 1 & -1 \\ 1 & -2 & 2 \end{pmatrix}$$

se obtiene al sustraer el segundo renglón del primero. En consecuencia, se infiere de inmediato que el rango de la matriz es 2. Por otro lado, $5 \neq 1 - 2$, de manera que no puede haber una solución del sistema de ecuaciones anterior.

Teorema 4.4. *Considere un sistema no homogéneo de ecuaciones lineales*

$$AX = B.$$

Suponga que existe al menos una solución X_0. Entonces el conjunto de soluciones es precisamente

$$X_0 + \mathrm{Ker}\, L_A.$$

En otras palabras, todas las soluciones son de la forma

$$X_0 + Y, \qquad \text{donde } Y \text{ es una solución de} \qquad AY = O.$$

Demostración. Sea $Y \in \mathrm{Ker}\, L_A$. Esto significa que $AY = O$. Entonces,

$$A(X_0 + Y) = AX_0 + AY = B + O = B,$$

de manera que $X_0 + \mathrm{Ker}\, L_A$ está contenido en el conjunto de soluciones. Recíprocamente, sea X cualquier solución de $AX = B$. Entonces,

$$A(X - X_0) = AX - AX_0 = B - B = O.$$

En consecuencia, $X = X_0 + (X - X_0)$, donde $X - X_0 = Y$ y $AY = O$. Esto prueba el teorema.

Cuando existe al menos una solución del sistema $AX = B$, entonces dim Ker L_A se conoce como **dimensión del conjunto de soluciones**. Es la dimensión del espacio de soluciones del sistema homogéneo.

Ejemplo 3. Encuentre la dimensión del conjunto de soluciones del siguiente sistema de ecuaciones y determine este conjunto de $\mathbf{R}^3$.

$$\begin{aligned} 2x + y + z &= 1, \\ y - z &= 0. \end{aligned}$$

Vemos por inspección que hay al menos una solución, a saber, $x = \frac{1}{2}, y = z = 0$. El rango de la matriz

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$

es igual a 2. Por tanto, la dimensión del conjunto de soluciones es 1. El espacio vectorial de las soluciones del sistema homogéneo tiene dimensión 1, y se encuentra fácilmente que una solución es

$$y = z = 1, \qquad x = -1.$$

En consecuencia, el conjunto de soluciones del sistema no homogéneo es el conjunto de todos los vectores

$$(\tfrac{1}{2}, 0, 0) + t(-1, 1, 1),$$

donde t varía sobre todos los números reales. Vemos que nuestro conjunto de soluciones es una línea recta.

Ejercicios IV, §4

1. Sea A un vector no nulo en $\mathbf{R}^n$. ¿Cuál es la dimensión del espacio de soluciones de la ecuación $A \cdot X = 0$?

2. ¿Cuál es la dimensión del subespacio de $\mathbf{R}^6$ perpendicular a los dos vectores $(1, 1, -2, 3, 4, 5)$ y $(0, 0, 1, 1, 0, 7)$?

3. Sea A un vector no nulo del espacio de n dimensiones. Sea P un punto en el espacio de n dimensiones. ¿Cuál es la dimensión del conjunto de soluciones de la ecuación
$$X \cdot A = P \cdot A?$$

4. ¿Cuál es la dimensión del espacio de soluciones de los siguientes sistemas de ecuaciones lineales? En cada caso, encuentre una base para el espacio de soluciones.
   (a) $\begin{aligned} 2x + y - z &= 0 \\ 2x + y + z &= 0 \end{aligned}$
   (b) $x - y + z = 0$
   (c) $\begin{aligned} 4x + 7y - \pi z &= 0 \\ 2x - y + z &= 0 \end{aligned}$
   (d) $\begin{aligned} x + y + z &= 0 \\ x - y &= 0 \\ y + z &= 0 \end{aligned}$

5. ¿Cuál es la dimensión del espacio de soluciones de los siguientes sistemas de ecuaciones lineales?
   (a) $\begin{aligned} 2x - 3y + z &= 0 \\ x + y - z &= 0 \end{aligned}$
   (b) $\begin{aligned} 2x + 7y &= 0 \\ x - 2y + z &= 0 \end{aligned}$
   (c) $\begin{aligned} 2x - 3y + z &= 0 \\ x + y - z &= 0 \\ 3x + 4y &= 0 \\ 5x + y + z &= 0 \end{aligned}$
   (d) $\begin{aligned} x + y + z &= 0 \\ 2x + 2y + 2z &= 0 \end{aligned}$

6. Sea $L: V \to W$ una aplicación lineal. Use un problema que aparece en el texto para probar que
$$\dim \operatorname{Im} L \leq \dim V$$

7. Sean A y B dos matrices que se pueden multiplicar entre sí, i.e., tales que AB existe. Pruebe que
$$\text{rango de } AB \leq \text{rango de } A \qquad \text{y} \qquad \text{rango de } AB \leq \text{rango de } B.$$

8. Sea U un subespacio de $\mathbf{R}^n$. Pruebe que $U^\perp$ también es un subespacio.

IV, §5. La matriz asociada con una aplicación lineal

A cada matriz A le hemos asociado una aplicación lineal L_A. Recíprocamente, dada una aplicación lineal
$$L: \mathbf{R}^n \to \mathbf{R}^m,$$
probaremos ahora que existe alguna matriz asociada A tal que $L = L_A$.

Sean $E^1, \ldots, E^n$ los vectores columna unitarios de $\mathbf{R}^n$. Para cada $j = 1, \ldots, n$, sea $L(E^j) = A^j$, donde A^j es un vector columna en $\mathbf{R}^m$. Por tanto,

$$L(E^1) = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} = a^1, \quad \ldots, \quad L(E^n) = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = a^n.$$

Entonces, para cada elemento X en $\mathbf{R}^n$, podemos escribir

$$X = x_1E^1 + \cdots + x_nE^n = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

y, por consiguiente,

$$\begin{aligned} L(X) &= x_1L(E^1) + \cdots + x_nL(E^n) \\ &= x_1A^1 + \cdots + x_nA^n \\ &= AX, \end{aligned}$$

donde A es la matriz cuyos vectores columna son $A^1, \ldots, A^n$. En consecuencia, $L = L_A$, lo cual prueba el teorema.

Observación. Al trabajar con $\mathbf{R}^n$ y $\mathbf{R}^m$ estamos en posibilidades de escribir vectores columna; es por eso que se pudo obtener en forma sencilla la matriz A que se acaba de derivar. Más adelante, en esta sección, trabajaremos con espacios vectoriales más generales, en términos de bases, y podremos escribir horizontalmente los vectores de coordenadas. La notación horizontal dará lugar a una transpuesta.

La matriz A anterior se denomina matriz **asociada** con la aplicación lineal L.

Tal como vimos al estudiar el espacio de columnas de A, podemos expresar las columnas de A en términos de los vectores unitarios:

$$(*) \qquad L(E^1) = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad \ldots, \quad L(E^n) = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

Ejemplo 1. Sea $F: \mathbf{R}^3 \to \mathbf{R}^2$ la proyección o, en otras palabras, la aplicación tal que $F^t(x_1, x_2, x_3) = t(x_1, x_2)$. Entonces la matriz asociada con F es

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Ejemplo 2. Sea $I: \mathbf{R}^n \to \mathbf{R}^n$ la matriz identidad. Entonces la matriz asociada con I es la matriz

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

que tiene componentes iguales a 1 en la diagonal y 0 en los demás lugares.

Ejemplo 3. Sea $L: \mathbf{R}^4 \to \mathbf{R}^2$ la aplicación lineal tal que

$$L(E^1) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad L(E^2) = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \quad L(E^3) = \begin{pmatrix} -5 \\ 4 \end{pmatrix}, \quad L(E^4) = \begin{pmatrix} 1 \\ 7 \end{pmatrix}.$$

Conforme a las relaciones (*), vemos que la matriz asociada con L es la matriz

$$\begin{pmatrix} 2 & 3 & -5 & 1 \\ 1 & -1 & 4 & 7 \end{pmatrix}.$$

Observación. Si en lugar de vectores columna usamos vectores renglón, entonces para hallar la matriz asociada daríamos origen a una *transpuesta*.

Sea V un espacio vectorial de dimensión n. Si escogemos alguna base $\{v_1, \ldots, v_n\}$ de V, entonces cada elemento de V se puede escribir en términos de coordenadas de la siguiente manera:

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Así, a cada elemento v de V le podemos asociar el **vector de coordenadas**

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Si

$$w = y_1 v_1 + \cdots + y_n v_n$$

de manera que Y es el vector de coordenadas de w, entonces

$$v + w = (x_1 + y_1)v_1 + \cdots + (x_n + y_n)v_n,$$

por lo que $X + Y$ es el vector de coordenadas de $v + w$. Sea c un número; entonces

$$cX = cx_1 v_1 + \cdots + cx_n v_n,$$

de manera que cX es el vector de coordenadas de cv. Por tanto, luego de escoger una base, podemos identificar V con $\mathbf{R}^n$ a través de los vectores de coordenadas.

Sea $L: V \to V$ una aplicación lineal. Entonces, después de escoger una base que nos dé una identificación de V con $\mathbf{R}^n$, podemos representar L mediante una matriz. Distintas elecciones de bases darán origen a menudo a distintas matrices asociadas. Algunas de las bases que se eligen suelen proporcionar matrices especialmente sencillas.

Ejemplo. Suponga que existe una base $\{v_1, \dots, v_n\}$ y que existen números $c_1, \dots, c_n$ tales que

$$Lv_i = c_i v_i \qquad \text{para} \qquad i = 1, \dots, n.$$

Entonces, con respecto a esta base, la matriz de L es la matriz diagonal

$$\begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & c_n \end{pmatrix}.$$

Si escogiéramos otra base, la matriz L podría no ser tan simple.

El principio general para hallar la matriz asociada de una aplicación lineal con respecto a una base se puede encontrar de la manera siguiente.

Sea $\{v_1, \dots, v_n\}$ la base dada de V. Entonces existen números c_{ij} tales que

$$\begin{aligned} Lv_1 &= c_{11}v_1 + \cdots + c_{1n}v_n \\ \vdots \quad & \quad \vdots \qquad\qquad \vdots \\ Lv_n &= c_{n1}v_1 + \cdots + c_{nn}v_n. \end{aligned}$$

¿Cuál es el efecto de L sobre el vector de coordenadas X de un elemento $v \in V$? Dicho elemento es de la forma

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Entonces

$$\begin{aligned} Lv &= \sum_{i=1}^{n} x_i L(v_i) \\ &= \sum_{i=1}^{n} x_i \sum_{j=1}^{n} c_{ij} v_j \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i c_{ij} v_j \\ &= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} x_i c_{ij} \right) v_j. \end{aligned}$$

En consecuencia, hallamos que

> *Si $C = (c_{ij})$ es la matriz tal que $L(v_i) = \sum_{j=1}^{n} c_{ij} v_j$ y X es el vector de coordenadas de v, entonces el vector de coordenadas de Lv es ${}^tCX$. En otras palabras, en vectores de coordenadas, L se representa mediante ${}^tC$ (transpuesta de C).*

Observemos que trabajamos con la *transpuesta* de C en lugar de la propia C. Esto se debe a que, al escribir LV_i como combinación lineal de $v_1, \dots, v_n$, la hemos escrito en forma horizontal, mientras que antes la escribimos verticalmente en términos de los vectores unitarios verticales $E^1, \dots, E^n$.

Ejemplo. Sea $L\colon V \to V$ una aplicación lineal. Sea $\{v_1, v_2, v_3\}$ una base de V tal que

$$L(v_1) = 2v_1 - v_2,$$
$$L(v_2) = v_1 + v_2 - 4v_3,$$
$$L(v_3) = 5v_1 + 4v_2 + 2v_3.$$

Entonces la matriz asociada con L sobre los vectores de coordenadas es la matriz

$$\begin{pmatrix} 2 & 1 & 5 \\ -1 & 1 & 4 \\ 0 & -4 & 2 \end{pmatrix},$$

que es la *transpuesta* de la matriz

$$\begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & -4 \\ 5 & 4 & 2 \end{pmatrix}.$$

Apéndice: Cambio de bases

Quizá el lector se pregunte cómo cambia la matriz que representa una aplicación lineal cuando cambiamos una base de V. Podemos encontrar la respuesta fácilmente de la manera siguiente. Primero veamos cómo cambian las coordenadas.

Sean $\{v_1, \dots, v_n\}$ una base y $\{w_1, \dots, w_n\}$ otra base de V.

Denótense con X las coordenadas de un vector con respecto a $\{v_1, \dots, v_n\}$ y denótense con Y las coordenadas del mismo vector con respecto a $\{w_1, \dots, w_n\}$.

¿En qué difieren X y Y? Ahora daremos la respuesta. Sea v el elemento de V que tiene coordenadas X, considerado como un vector columna. Así,

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Esto se asemeja a un producto interior y será conveniente usar la siguiente notación:

$$v = {}^tX \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = (x_1, \dots, x_n) \begin{pmatrix} v_1 \\ \vdots \\ vn \end{pmatrix},$$

donde ${}^tX$ es ahora una n-tupla renglón. En forma análoga,

$$v = {}^tY \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = (y_1, \dots, y_n) \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}.$$

Podemos expresar cada w_i como combinación lineal de los elementos de la base $v_1, \dots, v_n$, de manera que existe una matriz $B = (b_{ij})$ tal que, para cada i,

$$\boxed{w_i = \sum_{j=1}^{n} b_{ij} v_j \,.}$$

Sin embargo, podemos escribir estas relaciones en forma más eficiente en forma matricial

$$\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = B \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Por consiguiente, la relación para v en términos de los elementos de las bases se puede expresar en la siguiente forma:

$$v = {}^tY \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = {}^tYB \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \qquad \text{y también} \qquad v = {}^tX \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Por tanto, debemos tener ${}^tYB = {}^tX$. Tomar la transpuesta nos da la relación deseada

$$\boxed{X = {}^tBY\,.}$$

Note de nuevo la transpuesta. El cambio de coordenadas de una base a otra está dado en términos de la transpuesta de la matriz que expresa cada w_i como combinación lineal de $v_1, \dots, v_n$.

Observación. *La matriz B es invertible.*

Demostración. Hay varias maneras de ver esto. Por ejemplo, con los mismos argumentos que hemos dado, yendo de una base a otra, existe una matriz C tal que

$$Y = {}^tCX.$$

Esto es cierto para todas las n-tuplas coordenadas X y Y. Así, obtenemos

$$X = {}^tB{}^tCX = {}^t(CB)X.$$

para todas las n-tuplas X. En consecuencia, $CB = I$ es la matriz identidad. En forma análoga, $BC = I$, de manera que B es invertible.

Ahora sea $L: V \to V$ una aplicación lineal. Sea M la matriz que representa a L con respecto a la base $\{v_1, \dots, v_n\}$ y sea M' la matriz que representa a L con respecto a la base $\{w_1, \dots, w_n\}$. Por definición,

$$\text{las coordenadas de } L(v) \text{ son } \begin{cases} MX \text{ con respecto a } \{v_1, \dots, v_n\}, \\ M'Y \text{ con respecto a } \{w_1, \dots, w_n\}. \end{cases}$$

Por lo que acabamos de ver, tenemos necesariamente que

$$MX = {}^tBM'Y.$$

Sustituya $X = {}^tBY$ y multiplique por la izquierda ambos lados de la igualdad por ${}^tB^{-1}$. Entonces encontramos que

$${}^tB^{-1}M\,{}^tBY = M'Y.$$

Esto es cierto para todo Y. Si hacemos que $N = {}^tB$, entonces obtenemos la matriz M' en términos de M, a saber,

$$\boxed{M' = N^{-1}MN} \qquad \text{donde} \qquad N = {}^tB.$$

Por tanto, la matriz que representa a la aplicación lineal cambia mediante una transformación de semejanza. También podemos decir que M y M' son semejantes. En general, dos matrices M y M' son **semejantes** si existe una matriz invertible N tal que $M' = N^{-1}MN$.

En la práctica, no conviene escoger bases demasiado pronto. Para muchos problemas se debe seleccionar una base para la cual la matriz que representa a la aplicación lineal sea la más simple, y trabajar con esta base.

Ejemplo. Suponga que, con respecto a alguna base, la matriz M que representa a L es diagonal, digamos

$$M = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}.$$

Entonces la matriz que representa a L con respecto a otra base será de la forma

$$N^{-1}MN,$$

la que puede parecer un terrible revoltijo. Cambiar N en forma arbitraria corresponde a escoger una base arbitraria (por supuesto, N debe ser invertible). Más adelante, cuando estudiemos vectores propios, encontraremos condiciones en las cuales una matriz que representa una aplicación lineal es diagonal con respecto a una conveniente elección de base.

Ejercicios IV, §5

1. Encuentre la matriz asociada con las siguientes aplicaciones lineales.
   (a) $F: \mathbf{R}^4 \to \mathbf{R}^2$, dada por $F^t(x_1, x_2, x_3, x_4) = {}^t(x_1, x_2)$ (la proyección).
   (b) La proyección de $\mathbf{R}^4$ en $\mathbf{R}^3$.
   (c) $F: \mathbf{R}^2 \to \mathbf{R}^2$ dado por $F^t(x, y) = {}^t(3x, 3y)$.
   (d) $F: \mathbf{R}^n \to \mathbf{R}^n$ dado por $F(X) = 7X$.
   (e) $F: \mathbf{R}^n \to \mathbf{R}^n$ dado por $F(X) = -X$.
   (f) $F: \mathbf{R}^4 \to \mathbf{R}^4$ dado por $F^t(x_1, x_2, x_3, x_4) = {}^t(x_1, x_2, 0, 0)$.

2. Sea c un número y sea $L: \mathbf{R}^n \to \mathbf{R}^n$ la aplicación lineal tal que $L(X) = cX$. ¿Cuál es la matriz asociada con esta aplicación lineal?

3. Sea $F: \mathbf{R}^3 \to \mathbf{R}^2$ la aplicación lineal indicada. ¿Cuál es la matriz asociada de F?
   (a) $F(E^1) = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$, $F(E^2) = \begin{pmatrix} -4 \\ 2 \end{pmatrix}$, $F(E^3) = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.
   (b) $F\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3x_1 - 2x_2 + x_3 \\ 4x_1 - x_2 + 5x_3 \end{pmatrix}$.

4. Sea V un espacio de 3 dimensiones cuya base es $\{v_1, v_2, v_3\}$. Sea $F: V \to V$ la aplicación lineal que se indica. Encuentre la matriz de F con respecto a la base dada.
   (a) $F(v_1) = 3v_2 - v_3$,
   $F(v_2) = v_1 - 2v_2 + v_3$,
   $F(v_3) = -2v_1 + 4v_2 + 5v_3$.
   (b) $F(v_1) = 3v_1$, $F(v_2) = -7v_2$, $F(v_3) = 5v_3$.
   (c) $F(v_1) = -2v_1 + 7v_3$,
   $F(v_2) = -v_3$,
   $F(v_3) = v_1$.

5. En el texto dimos una descripción de una matriz asociada con una aplicación lineal de un espacio vectorial en sí mismo, con respecto a una base. En forma más general, sean V y W dos espacios vectoriales. Sean $\{v_1, \ldots, v_n\}$ una base de V y $\{w_1, \ldots, w_m\}$ una base de W. Sea $L: V \to W$ una aplicación lineal. Describa el lector cómo podría asociar una matriz con esta aplicación lineal, indicando el efecto sobre los vectores de coordenadas.

6. Sea $L: V \to V$ una aplicación lineal. Sea $v \in V$. Decimos que v es un vector propio para L si existe un número c tal que $L(v) = cv$. Suponga que V tiene una base $\{v_1, \ldots, v_n\}$ que consta de vectores propios, donde $L(v_i) = c_i v_i$ para $i = 1, \ldots, n$. ¿Cuál es la matriz que representa a L con respecto a esta base?

CAPÍTULO V

Composición y aplicaciones inversas

V, §1. Composición de aplicaciones lineales

Sean U, V y W conjuntos. Sean

$$F: U \to V \qquad \text{y} \qquad G: V \to W$$

aplicaciones. Entonces podemos formar la aplicación compuesta de U en W denotada con $G \circ F$. Por definición, es la aplicación tal que

$$(G \circ F)(u) = G(F(u)) \qquad \text{para todo } u \text{ en } U.$$

Ejemplo 1. Sea A una matriz de $m \times n$ y sea B una matriz de $q \times m$. Entonces podemos formar el producto BA. Sea

$$L_A: \mathbf{R}^n \to \mathbf{R}^m \quad \text{la aplicación lineal tal que} \quad L_A(X) = AX$$

y sea

$$L_B: \mathbf{R}^m \to \mathbf{R}^q \quad \text{la aplicación lineal tal que} \quad L_B(Y) = BY.$$

Entonces podemos formar la aplicación lineal compuesta $L_B \circ L_A$ tal que

$$(L_B \circ L_A)(X) = L_B(L_A(X)) = L_B(AX) = BAX.$$

Así tenemos que

$$\boxed{L_B \circ L_A = L_{BA}.}$$

Vemos que la composición de aplicaciones lineales corresponden a la multiplicación de matrices.

Ejemplo 2. Sea A una matriz de $m \times n$ y sea

$$L_A : \mathbf{R}^n \to \mathbf{R}^m$$

la aplicación lineal usual tal que $L_A(X) = AX$. Sea C un vector de $\mathbf{R}^m$ y sea

$$T_C : \mathbf{R}^m \to \mathbf{R}^m$$

la traslación determinada por C, esto es, $T_C(Y) = Y + C$. Entonces, la aplicación compuesta $T_C \circ L_A$ se obtiene al aplicar primero L_A a un vector X y luego al trasladar mediante C. Así,

$$T_C \circ L_A(X) = T_C(L_A(X)) = T_C(AX) = AX + C.$$

Ejemplo 3 Sea V un espacio vectorial y sea w un elemento de V. Sea

$$T_w : V \to V$$

la traslación determinada por w, esto es, la aplicación tal que $T_w(v) = v + w$. Entonces tenemos que

$$T_{w_1}(T_{w_2}(v)) = T_{w_1}(v + w_2) = v + w_2 + w_1.$$

Por tanto,

$$T_{w_1} \circ T_{w_2} = T_{w_1 + w_2}.$$

Podemos expresar este hecho diciendo que la composición de dos traslaciones de nuevo es una traslación. Por supuesto, la traslación T_w no es una aplicación lineal si $w \neq O$, debido a que

$$T_w(O) = O + w = w \neq O,$$

y sabemos que una aplicación lineal tiene que mandar al O en el O.

Ejemplo 4. Rotaciones. Sea θ un número y sea $A(\theta)$ la matriz

$$A(\theta) = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

Entonces $A(\theta)$ representa una rotación que podemos denotar con R_θ. La rotación compuesta $R_{\theta_1} \circ R_{\theta_2}$ se obtiene de la multiplicación de matrices y, para cualquier vector X de $\mathbf{R}^2$, tenemos que

$$R_{\theta_1} \circ R_{\theta_2}(X) = A(\theta_1)A(\theta_2)X.$$

Esta rotación compuesta es tan sólo la rotación determinada por la suma de los ángulos, a saber, $\theta_1 + \theta_2$. Esto corresponde a la fórmula $A(\theta_1)A(\theta_2) = A(\theta_1 + \theta_2)$.

El siguiente enunciado es una propiedad importante de las aplicaciones.

Sean U, V, W y S conjuntos. Sean

$$F: U \to V, \qquad G: V \to W \qquad y \qquad H: W \to S$$

aplicaciones. Entonces

$$H \circ (G \circ F) = (H \circ G) \circ F.$$

Demostración. De nuevo en este caso, la prueba es muy sencilla. Por definición, tenemos que, para cualquier elemento u de U:

$$\big(H \circ (G \circ F)\big)(u) = H\big((G \circ F)(u)\big) = H\big(G\big(F(u)\big)\big).$$

Por otro lado,

$$\big((H \circ G) \circ F)\big)(u) = \big(H \circ G)(F(u)\big) = H\big(G\big(F(u)\big)\big).$$

Por definición, esto significa que $(H \circ G) \circ F = H \circ (G \circ F)$.

Teorema 1.1. *Sean U, V y W espacios vectoriales. Sean las aplicaciones lineales*

$$F: U \to V \qquad y \qquad G: V \to W,$$

entonces la aplicación compuesta $G \circ F$ también es una aplicación lineal.

Demostración. Esto es muy fácil de probar. Sean u y v elementos de U. Como F es lineal, tenemos que $F(u + v) = F(u) + F(v)$. En consecuencia,

$$(G \circ F)(u + v) = G\big(F(u + v)\big) = G\big(F(u) + F(v)\big).$$

Como G es lineal, obtenemos

$$G\big(F(u) + F(v)\big) = G\big(F(u)\big) + G\big(F(v)\big).$$

Por tanto,

$$(G \circ F)(u + v) = (G \circ F)(u) + (G \circ F)(v).$$

Luego, sea c un número. Entonces

$$\begin{aligned}(G \circ F)(cu) &= G\big(F(cu)\big) \\ &= G\big(cF(u)\big) \qquad \text{(debido a que F es lineal)} \\ &= cG\big(F(u)\big) \qquad \text{(debido a que G es lineal).}\end{aligned}$$

Esto prueba que $G \circ F$ es una aplicación lineal.

El siguiente teorema establece que algunas de las reglas de la aritmética concernientes al producto y a la suma de números también se aplica a la composición y suma de aplicaciones lineales.

Teorema 1.2. *Sean U, V y W espacios vectoriales. Sea*

$$F: U \to V$$

una aplicación lineal y sean G y H dos aplicaciones lineales de V en W. Entonces,

$$(G+H)\circ F = G\circ F + H\circ F.$$

Si C es un número, entonces

$$(cG)\circ F = c(G\circ F).$$

Si $T\colon U\to V$ es una aplicación lineal de U en V, entonces

$$G\circ(F+T) = G\circ F + G\circ T.$$

Todas las pruebas son sencillas. Sólo probaremos el primer aserto y dejamos los demás como ejercicios.

Sea u un elemento de U. Tenemos que

$$\begin{aligned}((G+H)\circ F)(u) &= (G+H)(F(u)) = G(F(u)) + H(F(u))\\ &= (G\circ F)(u) + (H\circ F)(u).\end{aligned}$$

Por definición, se infiere que $(G+H)\circ F = G\circ F + H\circ F$.

Al igual que con las matrices, vemos que la composición y la adición de aplicaciones lineales se comportan como la multiplicación y la adición de números. Sin embargo, la misma observación que hicimos con respecto a las matrices se aplica en este caso. Primero, podemos no tener conmutatividad, y segundo, no tenemos "división", excepto, como se analizará en la siguiente sección para las inversas, cuando existan.

Ejemplo 5. Sea

$$F\colon \mathbf{R}^3\to\mathbf{R}^3$$

la aplicación lineal dada por

$$F(x,y,z) = (x,y,0)$$

y sea G la aplicación lineal dada por

$$G(x,y,z) = (x,z,0).$$

Entonces $(G\circ F)(x,y,z) = (x,0,0)$, pero $(F\circ G)(x,y,z) = (x,z,0)$.

Por otro lado, sea

$$L\colon V\to V$$

una aplicación lineal de un espacio vectorial en sí mismo. Podemos iterar L varias veces de manera que, como es usual, obtenemos

$$L^2 = L\circ L,\qquad L^3 = L\circ L\circ L,\qquad \text{y así sucesivamente.}$$

Convengamos además que

$$L^0 = I = \text{aplicación identidad}.$$

Por tanto L^k es la iteración de L consigo misma k veces. Para dichas potencias de L sí tenemos conmutatividad, a saber,

$$L^{r+s} = L^r\circ L^s = L^s\circ L^r.$$

Ejercicios V, §1

1. Sean A y B dos matrices de $m \times n$. Suponga que
$$AX = BX$$
para todas las n-tuplas X. Muestre que $A = B$.

2. Sean F y L aplicaciones lineales de V en sí mismo. Suponga que F y L conmutan, esto es, que $F \circ L = L \circ F$. Pruebe las conocidas reglas siguientes:
$$(F + L)^2 = F^2 + 2F \circ L + L^2,$$
$$(F - L)^2 = F^2 - 2F \circ L + L^2,$$
$$(F + L) \circ (F - L) = F^2 - L^2.$$

3. Pruebe la conocida regla para una aplicación lineal $F: V \to V$:
$$(I - F) \circ (I + F + \cdots + F^r) = I - F^{r+1}.$$

4. Sean V un espacio vectorial y $T: V \to V$ una aplicación lineal tal que $T^2 = I$. Defina $P = \frac{1}{2}(I + T)$ y $Q = \frac{1}{2}(I - T)$.
   (a) Demuestre que $P^2 = P$ y que $Q^2 = Q$.
   (b) Demuestre que $P + Q = I$.
   (c) Demuestre que $\operatorname{Ker} P = \operatorname{Im} Q$ y que $\operatorname{Im} P = \operatorname{Ker} Q$.

5. Sea $P: V \to V$ una aplicación lineal tal que $P^2 = P$. Defina $Q = I - P$.
   (a) Demuestre que $Q^2 = Q$.
   (b) Demuestre que $\operatorname{Im} P = \operatorname{Ker} Q$ y que $\operatorname{Ker} P = \operatorname{Im} Q$.
   Una aplicación lineal P tal que $P^2 = P$ se conoce como **proyección** y generaliza la noción de proyección en el sentido usual.

6. Sea $P: V \to V$ una **proyección**, esto es, una aplicación lineal tal que $P^2 = P$.
   (a) Demuestre que $V = \operatorname{Ker} P + \operatorname{Im} P$.
   (b) Demuestre que la intersección de $\operatorname{Im} P$ con $\operatorname{Ker} P$ es $\{O\}$. En otras palabras, si $v \in \operatorname{Im} P$ y $v \in \operatorname{Ker} P$, entonces $v = O$.

Sea V un espacio vectorial y sean U y W subespacios. Se dice que V es una **suma directa** de U y W si se satisfacen las siguientes condiciones:
$$V = U + W \qquad \text{y} \qquad U \cap W = \{O\}.$$
En el ejercicio 6, el lector acaba de probar que V es la **suma directa** de $\operatorname{Ker} P$ e $\operatorname{Im} P$.

7. Sea V la suma directa de los subespacios U y W. Sea $v \in V$ y suponga que hemos expresado v como una suma $v = u + w$ con $u \in U$ y $w \in W$. Demuestre que u y w están determinados en forma única por v. Esto es, si $v = u_1 + w_1$ con $u_1 \in U$ y $w_1 \in W$, entonces $u = u_1$ y $w = w_1$.

8. Sean U y W dos espacios vectoriales, y sea $V = U \times W$ el conjunto de todas las parejas (u, w) con $u \in U$ y $w \in W$. Entonces V es un espacio vectorial como el que se describe en el Ejercicio 15 del Capítulo IV, sección §3. Sea
$$P: V \to V$$
la aplicación tal que $P(u, w) = (u, 0)$. Muestre que P es una proyección.

Si el lector identifica U con el conjunto de todos los elementos $(u, 0)$, donde $u \in U$, e identifica W con el conjunto de todos los elementos $(0, w)$ donde $w \in W$, entonces V es la suma directa de U y W.

Por ejemplo, sea $n = r + s$, donde r y s son enteros positivos. Entonces

$$\mathbf{R}^n = \mathbf{R}^r \times \mathbf{R}^s.$$

Observe que $\mathbf{R}^n$ es el conjunto de todas las n-tuplas de números reales, que se pueden considerar como

$$(x_1, \ldots, x_r, y_1, \ldots, y_s), \qquad \text{donde} \qquad x_i, y_j \in \mathbf{R}.$$

Por tanto, $\mathbf{R}^n$ se puede considerar como la suma directa de $\mathbf{R}^r$ y $\mathbf{R}^s$. La proyección de $\mathbf{R}^n$ sobre las primeras r componentes está dada por la aplicación P tal que

$$P(x_1, \ldots, x_r, y_1, \ldots, y_s) = (x_1, \ldots, x_r, 0, \ldots, 0).$$

Esta aplicación es una aplicación lineal y $P^2 = P$. Algunas veces, a la aplicación tal que

$$P(x_1, \ldots, x_n) = (x_1, \ldots, x_r)$$

también se le llama proyección, aunque aplica $\mathbf{R}^n$ en $\mathbf{R}^r$, de manera que no podemos aplicar nuevamente P a $(x_1, \ldots, x_r)$, puesto que P está definida en todo $\mathbf{R}^n$.

Observe que también se puede considerar la proyección sobre el segundo conjunto de coordenadas, esto es, hacemos

$$Q(x_1, \ldots, x_r, y_1, \ldots, y_s) = (0, \ldots, 0, y_1, \ldots, y_s).$$

Ésta se llama proyección sobre el segundo factor de $\mathbf{R}^n$, considerado como $\mathbf{R}^r \times \mathbf{R}^s$.

9. Sean A y B dos matrices tales que está definido el producto AB. Pruebe que

$$\text{rango de } AB \leq \text{rango de } A \qquad \text{y} \qquad \text{rango de } AB \leq \text{rango } B.$$

[*Sugerencia:* Considere las aplicaciones lineales L_{AB}, L_A y L_B.]

En efecto, defina el rango de una aplicación lineal L como $\dim \operatorname{Im} L$. Si $L: V \to W$ y $F: W \to U$ son aplicaciones lineales de espacios de dimensión finita, demuestre que

$$\text{rango de } F \circ L \leq \text{rango de } F \qquad \text{y} \qquad \text{rango de } F \circ L \leq \text{rango de } L.$$

10. Sea A una matriz de $m \times n$ y sea C un vector de $\mathbf{R}^m$. Sea T_C la traslación determinada por C, como en el Ejemplo 2. Escriba con detalle las fórmulas para

$$L_A \circ T_C(X) \qquad \text{y} \qquad T_C \circ L_A(X),$$

donde X está en $\mathbf{R}^n$. Dé un ejemplo de A y C tal que

$$L_A \circ T_C \neq T_C \circ L_A.$$

V, §2. Inversas

Sea

$$F: V \to W$$

una aplicación (usualmente se considera lineal). Decimos que F tiene una **inversa** si existe una aplicación

$$G: W \to V$$

tal que

$$G \circ F = I_V \quad \text{y} \quad F \circ G = I_W.$$

Con esto queremos decir que las aplicaciones compuestas $G \circ F$ y $F \circ G$ son las aplicaciones identidad de V y W, respectivamente. Si F tiene una inversa, también decimos que F es **invertible**.

Ejemplo 1. La aplicación inversa de la traslación T_u es la traslación T_{-u}, debido a que

$$T_{-u} \circ T_u(v) = T_{-u}(v + u) = v + u - u = v.$$

Así,

$$T_{-u} \circ T_u = I.$$

En forma análoga, $T_u \circ T_{-u} = I$.

Ejemplo 2. Sea A una matriz cuadrada de $n \times n$, y sea

$$L_A: \mathbf{R}^n \to \mathbf{R}^n$$

la aplicación lineal usual tal que $L_A(X) = AX$. Suponga que A tiene una matriz inversa A^{-1}, de manera que $AA^{-1} = A^{-1}A = I$. Entonces la fórmula

$$L_A \circ L_B = L_{AB}$$

de la sección anterior muestra que

$$L_A \circ L_{A^{-1}} = L_I = I.$$

En consecuencia, L_A tiene una aplicación inversa, que es precisamente la multiplicación por A^{-1}.

Teorema 2.1. *Sea $F: U \to V$ una aplicación lineal y supongamos que esta aplicación tiene una aplicación inversa $G: V \to U$. Entonces G es una aplicación lineal.*

Demostración. Sean $v_1, v_2 \in V$. Primero debemos demostrar que

$$G(v_1 + v_2) = G(v_1) + G(v_2).$$

Sean $u_1 = G(v_1)$ y $u_2 = G(v_2)$. Por definición, esto significa que

$$F(u_1) = v_1 \quad \text{y} \quad F(u_2) = v_2.$$

Como F es lineal, hallamos que

$$F(u_1 + u_2) = F(u_1) + F(u_2) = v_1 + v_2.$$

Por definición de aplicación inversa, esto significa que $G(v_1 + v_2) = u_1 + u_2$, con lo que se prueba lo que se quería. Dejamos como ejercicio la comprobación de que $G(cv) = cG(v)$ (Ejercicio 2).

Ejemplo 3. Sea $L: V \to V$ una aplicación lineal tal que $L^2 = O$. Entonces $I + L$ es invertible, debido a que

$$(I + L)(I - L) = I^2 - L^2 = I,$$

y en forma análoga, por otra parte, $(I - L)(I + L) = I$. Así tenemos

$$(I + L)^{-1} = I - L.$$

Ahora expresaremos las aplicaciones inversas con una terminología un poco diferente.

Sea

$$F: V \to W$$

una aplicación. Decimos que F es **inyectiva** (**uno a uno** en terminología antigua) si, dados elementos v_1 y v_2 de V tales que $v_1 \neq v_2$, entonces $F(v_1) \neq F(v_2)$.

Nos interesa especialmente esta noción para las aplicaciones lineales.

Ejemplo 4. Suponga que F es una aplicación lineal cuyo núcleo no es $\{O\}$. Entonces existe un elemento $v \neq O$ en el núcleo, y tenemos

$$F(O) = F(v) = O.$$

Por tanto, F no es inyectiva. Ahora probaremos el recíproco.

Teorema 2.2. *Una aplicación lineal $F: V \to W$ es inyectiva si, y sólo si, su núcleo es $\{O\}$.*

Demostración. Ya hemos demostrado una implicación. Recíprocamente, supongamos que el núcleo es $\{O\}$. Debemos probar que F es inyectiva. Sean $v_1 \neq v_2$ elementos de V distintos entre sí. Debemos mostrar que $F(v_1) \neq F(v_2)$. Pero

$$F(v_1) - F(v_2) = F(v_1 - v_2) \quad \text{debido a que } F \text{ es lineal.}$$

Como el núcleo de F es $\{O\}$ y $v_1 - v_2 \neq O$, se infiere que $F(v_1 - v_2) \neq O$. En consecuencia, $F(v_1) - F(v_2) \neq O$, por lo que $F(v_1) \neq F(v_2)$. Esto prueba el teorema.

Sea $F: V \to W$ una aplicación. Si la imagen de F es todo W, entonces decimos que F es **suprayectiva**. Las dos nociones de aplicación inyectiva o suprayectiva se combinan para dar un criterio básico para que F tenga una inversa.

Teorema 2.3. *Una aplicación $F: V \to W$ tiene una inversa si, y sólo si, es inyectiva y suprayectiva.*

Demostración. Supogamos que F es inyectiva y suprayectiva. Dado un elemento w de W, existe un elemento v de V tal que $F(v) = w$ (debido a que F es suprayectiva). Existe sólo uno de tales elementos v (debido a que F es inyectiva). Así, podemos definir

$$G(w) = \text{único elemento } v \text{ tal que } F(v) = w.$$

Por la manera en que hemos definido G, resulta claro que

$$G(F(v)) = v \qquad \text{y} \qquad F(G(w)) = w.$$

Por tanto, G es la aplicación inversa de F.

Recíprocamente, suponga que F tiene una aplicación inversa G. Sean v_1 y v_2 elementos de V tales que $F(v_1) = F(v_2)$. Al aplicar G se obtiene

$$v_1 = G \circ F(v_1) = G \circ F(v_2) = v_2,$$

de manera que F es inyectiva. Luego consideramos un elemento w de W; la ecuación

$$w = F(G(w))$$

muestra que $w = F(v)$ para algún v, a saber, $v = G(w)$, por lo que F es suprayectiva. Esto prueba el teorema.

En el caso de las aplicaciones lineales, tenemos ciertos criterios para verificar la inyectividad o la suprayectividad, lo que nos permite verificar menos condiciones cuando deseamos probar que una aplicación lineal es invertible.

Teorema 2.4. *Sea $F: V \to W$ una aplicación lineal. Suponga que*

$$\dim V = \dim W.$$

(i) *Si* $\operatorname{Ker} F = \{O\}$, *entonces F es invertible.*
(ii) *Si F es suprayectiva, entonces F es invertible.*

Demostración. Supongamos primero que $\operatorname{Ker} F = \{O\}$. Entonces F es inyectiva, por el Teorema 2.2. Pero

$$\dim V = \dim \operatorname{Ker} F + \dim \operatorname{Im} F,$$

de manera que $\dim V = \dim \operatorname{Im} F$, y la imagen de F es un subespacio de W que tiene la misma dimensión que W. Por tanto, $\operatorname{Im} F = W$, por el Teorema 5.6 del Capítulo III. En consecuencia, F es suprayectiva. Esto prueba (i), al usar el Teorema 2.3.

Se deja como ejercicio la prueba de (ii).

Ejemplo 5. Sea $F: \mathbf{R}^2 \to \mathbf{R}^2$ una aplicación lineal tal que

$$F(x, y) = (3x - y, 4x + 2y).$$

Deseamos mostrar que F tiene inversa. Primero observamos que el núcleo de F es $\{O\}$, debido a que, si

$$3x - \ y = 0,$$
$$4x + 2y = 0,$$

entonces podemos resolver para x y y de la manera usual: multipliquemos la primera ecuación por 2 y sumémosla a la segunda ecuación. Encontramos que $10x = 0$, por lo que $x = 0$ y , así, $y = 0$ debido a que $y = 3x$. En consecuencia, F es inyectiva, ya que su núcleo es $\{O\}$.

Por tanto, F es invertible, por el Teorema 2.4(i).

Una aplicación lineal $F: U \to V$ que tiene una inversa $G: V \to U$ (también decimos que es **invertible**) se conoce como **isomorfismo**.

Ejemplo 6. Sea V un espacio vectorial de dimensión n. Sea

$$\{v_1, \ldots, v_n\}$$

una base de V. Sea

$$L: \mathbf{R}^n \to V$$

la aplicación tal que

$$L(x_1, \ldots, x_n) = x_1 v_1 + \cdots + x_n v_n.$$

Entonces L es un isomorfismo.

Demostración. El núcleo de L es $\{O\}$, debido a que, si

$$x_1 v_1 + \cdots + x_n v_n = O,$$

entonces todo $x_i = 0$ (puesto que $v_1, \ldots, v_n$ son linealmente independientes). La imagen de L es todo V, ya que $v_1, \ldots, v_n$ generan a V. Por el Teorema 2.4, se infiere que L es un isomorfismo.

Ejercicios V, §2

1. Sea R_θ la rotación levógira en un ángulo θ. ¿Cómo expresaría de una manera sencilla la inversa R_θ^{-1} como R_θ para algún φ? Si

$$A(\theta) = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

es la matriz asociada con R_θ, ¿cuál es la matriz asociada con R_θ^{-1}?

2. (a) Termine la demostración del Teorema 2.1.
   (b) Dé la demostración del Teorema 2.4(ii).

3. Sean F y G aplicaciones lineales invertibles de un espacio vectorial V en sí mismo. Demuestre que

$$(F \circ G)^{-1} = G^{-1} \circ F^{-1}.$$

4. Sea $L: \mathbf{R}^2 \to \mathbf{R}^2$ la aplicación lineal definida por

$$L(x, y) = (x + y, x - y).$$

Demuestre que L es invertible.

5. Sea $L: \mathbf{R}^2 \to \mathbf{R}^2$ la aplicación lineal definida por

$$L(x, y) = (2x + y, 3x - 5y).$$

Demuestre que L es invertible.

6. Sea $L: \mathbf{R}^3 \to \mathbf{R}^3$ cada una de las aplicaciones lineales indicadas. Demuestre que L es invertible en cada caso.
   (a) $L(x, y, z) = (x - y, x + z, x + y + 3z)$
   (b) $L(x, y, z) = (2x - y + z, x + y, 3x + y + z)$.

7. Sea $L: V \to V$ una aplicación lineal tal que $L^2 = O$. Demuestre que $I - L$ es invertible. (I es la aplicación identidad sobre V.)

8. Sea $L: V \to V$ una aplicación lineal tal que $L^2 + 2L + I = O$. Demuestre que L es invertible.

9. Sea $L: V \to V$ una aplicación lineal tal que $L^3 = O$. Demuestre que $I - L$ es invertible.

10. Sea $L: V \to V$ una aplicación lineal tal que $L^r = O$. Demuestre que $I - L$ es invertible.

11. Sea V un espacio vectorial de dos dimensiones y sea $L: V \to V$ una aplicación lineal tal que $L^2 = O$, pero $L \neq O$. Sea v un elemento de V tal que $L(v) \neq O$. Sea $w = L(v)$. Pruebe que $\{v, w\}$ es una base de V.

12. Sea V el conjunto de todas las sucesiones infinitas de números reales

$$(x_1, x_2, x_3, \ldots).$$

Éste podría llamarse espacio de dimensión infinita. La adición y la multiplicación por números se define componente a componente, de manera que V es un espacio vectorial. Defina la aplicación $F: V \to V$ como

$$F(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, x_3, \ldots).$$

Por razones obvias, F se conoce como el **operador de traslación** y F es lineal.
   (a) ¿Es inyectiva F? ¿Cuál es el núcleo de F?
   (b) ¿Es suprayectiva F?
   (c) Demuestre que hay una aplicación lineal $G: V \to V$ tal que $G \circ F = I$.
   (d) La aplicación G que aparece en (c) ¿tiene la propiedad $F \circ G = I$?

13. Sea V un espacio vectorial y sean U y W dos subespacios. Suponga que V es la suma directa de U y W, esto es,

$$V = U + W \qquad \text{y} \qquad U \cap W = \{O\}.$$

Sea $L: U \times W \to V$ la aplicación tal que

$$L(u, w) = u + w.$$

Demuestre que L es una aplicación lineal biyectiva. (Que una aplicación lineal sea biyectiva significa que es inyectiva y suprayectiva.)

CAPÍTULO VI

Productos escalares y ortogonalidad

VI, §1. Productos escalares

Sea V un espacio vectorial. Un **producto escalar** sobre V es una asociación que a cualquier par de elementos (v, w) de V le asocia un número, denotado con $\langle v, w\rangle$, que satisface las siguientes propiedades:

PE 1. *Tenemos que* $\langle v, w\rangle = \langle w, v\rangle$ *para todos los* v *y* w *en* V.

PE 2. *Si* u, v *y* w *son elementos de* V, *entonces*

$$\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle.$$

PE 3. *Si* x *es un número, entonces*

$$\langle xu, v\rangle = x\langle u, v\rangle = \langle u, xv\rangle.$$

También supondremos que el producto escalar satisface la siguiente condición:

PE 4. *Para todo* v *en* V *tenemos que* $\langle v, v\rangle \geq 0$ *y* $\langle v, v\rangle > 0$ *si* $v \neq O$.

Un producto escalar que satisface esta condición se conoce como **definitivamente positivo.**

En lo que resta de esta sección supondremos que V *es un espacio vectorial con un producto escalar definitivamente positivo.*

Ejemplo 1. Sea $V = \mathbf{R}^n$, y defina

$$\langle X, Y\rangle = X \cdot Y$$

para los elementos X y Y de $\mathbf{R}^n$. Entonces éste es un producto escalar definitivamente positivo.

Ejemplo 2. Sea V el espacio de funciones continuas con valores reales en el intervalo $[-\pi, \pi]$. Si f y g están en V, definimos

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(t)g(t)\, dt.$$

Algunas propiedades sencillas de la integral muestran que éste es un producto escalar, que es, en efecto, definitivamente positivo.

En cálculo estudiamos el segundo ejemplo, que da origen a la teoría de las series de Fourier. En este libro sólo estudiamos las propiedades generales de los productos escalares y sus aplicaciones a espacios euclidianos. Se emplea la notación $\langle\ ,\ \rangle$ porque, al trabajar con espacios vectoriales de funciones, el usar un punto $f \cdot g$ se puede confundir con el producto ordinario de funciones.

Igual que en el caso de producto interior, definimos los elementos v, w de V como **ortogonales** o **perpendiculares entre sí**, y escribimos $v \perp w$, si $\langle v, w\rangle = 0$. Si S es un subconjunto de V, denotamos con $S^{\perp}$ el conjunto de todos los elementos w de V que son perpendiculares a todos los elementos de S, i.e., tales que $\langle w, v\rangle = 0$ para todo v en S. Luego, usando **PE 1, PE 2** y **PE 3** se verifica fácilmente que $S^{\perp}$ es un subespacio de V, conocido como el **espacio ortogonal** de S. Si w es perpendicular a S, escribimos también $w \perp S$. Sea U el subespacio de V generado por los elementos de S. Si w es perpendicular a S y si v_1 y v_2 están en S, entonces

$$\langle w, v_1 + v_2\rangle = \langle w, v_1\rangle + \langle w, v_2\rangle = 0.$$

Si c es un número, entonces

$$\langle w, cv_1\rangle = c\langle w, v_1\rangle = 0.$$

En consecuencia, w es perpendicular a las combinaciones lineales de elementos de S y, por tanto, w es perpendicular a U.

Ejemplo 3. Sea (a_{ij}) una matriz de $m \times n$ y sean $A_1, \dots, A_m$ sus vectores renglón. Sea $X = (x_1, \dots, x_n)$ como de costumbre. El sistema de ecuaciones lineales homogéneas

$$(**) \qquad \begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ \vdots \qquad \vdots \qquad & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

también se puede escribir en forma abreviada usando el producto interior de la forma siguiente:

$$A_1 \cdot X = 0, \quad \ldots, \quad A_m \cdot X = 0.$$

El conjunto de soluciones X de este sistema homogéneo es, por consiguiente, el conjunto de todos los vectores perpendiculares a $A_1, \ldots, A_m$. Por tanto, es el subespacio de $\mathbf{R}^n$ que es el subespacio ortogonal al espacio generado por $A_1, \ldots A_m$. Si U es el espacio de soluciones, y si W denota el espacio generado por $A_1, \ldots A_m$, tenemos

$$U = W^{\perp}.$$

Decimos que dim U es la **dimensión del espacio de soluciones del sistema de ecuaciones lineales.**

Igual que en el Capítulo I, definimos la **longitud** o **norma** de un elemento $v \in V$ como

$$||v|| = \sqrt{\langle v, v \rangle}.$$

Si c es cualquier número, entonces de inmediato obtenemos

$$||cv|| = |c| \, ||v||,$$

debido a que

$$||cv|| = \sqrt{\langle cv, cv \rangle} = \sqrt{c^2 \langle v, v \rangle} = |c| \, ||v||.$$

Así, vemos que el mismo tipo de argumentos usados en el Capítulo I se puede emplear en esta parte. De hecho, cualquier argumento dado en el Capítulo I que no use coordenadas se aplica a nuestra situación más general. Al avanzar en el tema veremos más ejemplos.

Igual que antes, decimos que un elemento $v \in V$ es un **vector unitario** si $||v|| = 1$. Si $v \in V$ y $v \neq O$, entonces $v/||v||$ es un vector unitario.

Las siguientes dos identidades se infieren en forma directa de la definición de longitud.

El teorema de Pitágoras. *Si v, w son perpendiculares entre sí, entonces*

$$||v + w||^2 = ||v||^2 + ||w||^2.$$

La ley del paralelogramo. *Para cualesquiera v y w tenemos que*

$$||v + w||^2 + ||v - w||^2 = 2||v||^2 + 2||w||^2.$$

Las pruebas son triviales. Damos la primera y dejamos la segunda como ejercicio. Para la primera, tenemos que

$$\begin{aligned} ||v + w||^2 = \langle v + w, v + w \rangle &= \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle \\ &= ||v||^2 + ||w||^2. \end{aligned}$$

Sea w un elemento de V tal que $||w|| \neq 0$. Para cualquier v existe un único número c tal que $v - cw$ es perpendicular a w. Ciertamente, para que $v - cw$ sea perpendicular a w, debemos tener

$$\langle v - cw, w \rangle = 0,$$

de ahí que $\langle v, w\rangle - \langle cw, w\rangle = 0$, y $\langle v, w\rangle = c\langle w, w\rangle$. Por tanto,

$$c = \frac{\langle v, w\rangle}{\langle w, w\rangle}.$$

Recíprocamente, al hacer que c tome este valor se muestra que $v - cw$ es perpendicular a w. Decimos que c es la **componente de v a lo largo de w**.

En particular, si w es un vector unitario, entonces la componente de v a lo largo de w es simplemente

$$c = \langle v, w\rangle.$$

Ejemplo 4. Sea $V = \mathbf{R}^n$ con el producto escalar usual, i.e., el producto interior. Si E_i es el i-ésimo vector unitario y $X = (x_1, \ldots, x_n)$, entonces la componente de X a lo largo de E_i es sencillamente

$$X \cdot E_i = x_i,$$

esto es, la i-ésima componente de X.

Ejemplo 5. Sea V el espacio de funciones continuas sobre $[-\pi, \pi]$. Sea f la función dada por $f(x) = \operatorname{sen} kx$, donde k es algún entero > 0. Entonces,

$$\|f\| = \sqrt{\langle f, f\rangle} = \left(\int_{-\pi}^{\pi} \operatorname{sen}^2 kx\, dx\right)^{1/2} = \sqrt{\pi}.$$

Si g es cualquier función continua sobre $[-\pi, \pi]$, entonces la componente de g a lo largo de f también se conoce como **coeficiente de Fourier de g con respecto a f**, y es igual a

$$\frac{\langle g, f\rangle}{\langle f, f, \rangle} = \frac{1}{\pi}\int_{-\pi}^{\pi} g(x) \operatorname{sen} kx\, dx.$$

Como en el caso del espacio de n dimensiones, definimos la **proyección de v a lo largo de w** como el vector cw, a causa de nuestra representación usual:

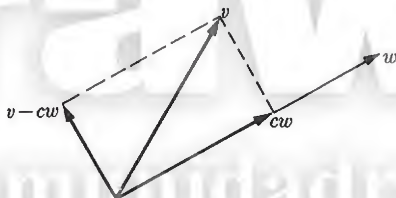


Figura 1

Ahora se pueden usar exactamente los mismos argumentos que dimos en el Capítulo I para obtener la **desigualdad de Schwarz**, a saber,

Teorema 1.1. *Para todos los v y $w \in V$ tenemos*

$$|\langle v, w\rangle| \leq \|v\|\,\|w\|.$$

Demostración. Si $w = O$, entonces ambos lados son iguales a 0 y nuestra desigualdad es obvia. Ahora bien, suponga que $w \neq O$. Sea c la componente de v a lo largo de w. Escribimos

$$v = v - cw + cw.$$

Entonces $v - cw$ es perpendicular a cw, de manera que, por Pitágoras,

$$\begin{aligned}\|v\|^2 &= \|v - cw\|^2 + \|cw\|^2 \\ &= \|v - cw\|^2 + |c|^2\,\|w\|^2.\end{aligned}$$

Por consiguiente, $|c|^2\,\|w\|^2 \leq \|v\|^2$ y al extraer raíces cuadradas se obtiene

$$|c|\,\|w\| \leq \|v\|.$$

Pero $c = \langle v, w\rangle/\|w\|^2$. Entonces un factor $\|w\|$ se cancela y, al multiplicar ambos miembros de la desigualdad por $\|w\|$, se obtiene

$$|\langle v, w\rangle| \leq \|v\|\,\|w\|,$$

lo que prueba el teorema.

Teorema 1.2. *Si v y $w \in V$, entonces*

$$\|v + w\| \leq \|v\| + \|w\|.$$

Demostración. Es exactamente la misma que la del teorema análogo que aparece en el Capítulo I, sección §4.

Sean $v_1, \ldots, v_n$ elementos no nulos de V y mutuamente perpendiculares, esto es, $\langle v_i, v_j\rangle = 0$ si $i \neq j$. Sea c_i la componente de v a lo largo de v_i. Entonces

$$v - c_1v_1 - \cdots - c_nv_n$$

es perpendicular a $v_1, \ldots, v_n$. Para ver esto, todo lo que tenemos que hacer es considerar el producto con v_j para cualquier j. Todos los términos que incluyen $\langle v_i, v_j\rangle$ darán 0 si $i \neq j$, y tendremos dos términos restantes

$$\langle v, v_j\rangle - c_j\langle v_j, v_j\rangle$$

que se cancelan. Así, al restar combinaciones lineales como antes, se ortogonaliza v con respecto a $v_1, \ldots, v_n$. El siguiente teorema muestra que $c_1v_1 + \cdots + c_nv_n$ da la mejor aproximación a v como combinación lineal de $v_1, \ldots, v_n$.

Teorema 1.3. *Sean $v_1, \ldots, v_n$ vectores mutuamente perpendiculares y tales que $\|v_i\| \neq 0$ para todo i. Sea v un elemento de V y sea c_i la componente de v a lo largo v_i. Sean $a_1, \ldots, a_n$ números. Entonces*

$$\|v - \textstyle\sum_{k=1}^{n} c_kv_k\| \leq \|v - \sum_{k=1}^{n} a_kv_k\|.$$

Demostración. Sabemos que

$$v - \sum_{k=1}^{n} c_k v_k$$

es perpendicular a cada v_i, $i = 1, \ldots, n$. En consecuencia, es perpendicular a cualquier combinación lineal de $v_1, \ldots, v_n$. Ahora tenemos:

$$\begin{aligned}\|v - \sum a_k v_k\|^2 &= \|v - \sum c_k v_k + \sum (c_k - a_k) v_k\|^2 \\ &= \|v - \sum c_k v_k\|^2 + \|\sum (c_k - a_k) v_k\|^2\end{aligned}$$

por el teorema de Pitágoras. Esto prueba que

$$\|v - \sum c_k v_k\|^2 \leq \|v - \sum a_k v_k\|^2,$$

con lo que nuestro teorema está probado.

Ejemplo 6. Considere el espacio vectorial V de todas las funciones continuas sobre el intervalo $[0, 2\pi]$. Sea

$$g_k(x) = \cos kx, \qquad \text{para} \qquad k = 0, 1, 2, \ldots.$$

Usemos el producto escalar

$$\langle f, g \rangle = \int_0^{2\pi} f(x) g(x)\, dx.$$

Entonces se verifica fácilmente que

$$\|g_0\| = \sqrt{2\pi} \qquad \text{y} \qquad \|g_k\| = \sqrt{\pi} \qquad \text{para} \qquad k > 0.$$

El coeficiente de Fourier de f con respecto a g_k es

$$c_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx, dx, \qquad \text{para} \qquad k > 0.$$

Si tomamos $v_k = g_k$ para $k = 1, \ldots, n$, entonces el Teorema 1.3 nos dice que la combinación lineal

$$c_0 + c_1 \cos x + c_2 \cos 2x + \cdots + c_n \cos nx$$

da la mejor aproximación a la función f entre todas las posibles combinaciones lineales

$$a_0 + a_1 \cos x + \cdots + a_n \cos nx$$

donde $a_0, a_1, \ldots, a_n$ son números reales arbitrarios. Dicha suma se conoce como **suma parcial** de la serie de Fourier.

En forma análoga, podríamos tomar combinaciones lineales de las funciones sen kx. Esto conduce a la teoría de las series de Fourier. En este libro no vamos a profundizar más. Solamente quisimos señalar la analogía y la utilidad del lenguaje geométrico y el formalismo al trabajar con estos objetos.

El siguiente teorema se conoce como la **desigualdad de Bessel**.

Teorema 1.4. *Si $v_1, \ldots, v_n$ son vectores unitarios mutuamente perpendiculares, y si c_i es el coeficiente de Fourier de v con respecto a v_i, entonces*

$$\sum_{i=1}^{n} c_i^2 \leq \|v\|^2.$$

Demostración. Tenemos que

$$\begin{aligned} 0 &\leq \langle v - \sum c_i v_i, v - \sum c_i v_i \rangle \\ &= \langle v, v \rangle - \sum 2c_i \langle v, v_i \rangle + \sum c_i^2 \\ &= \langle v, v \rangle - \sum c_i^2. \end{aligned}$$

A partir de esto se infiere nuestra desigualdad.

Ejercicios VI, §1

1. Sea V un espacio vectorial con un producto escalar definitivamente positivo. Sean $v_1, \ldots, v_r$ elementos de V, distintos de cero, que son mutuamente perpendiculares, lo que significa que $\langle v_i, v_j \rangle = 0$ si $i \neq j$. Demuestre que $v_1, \ldots, v_r$ son linealmente independientes.

El siguiente ejercicio proporciona un ejemplo importante de un producto escalar.

2. Sea A una matriz simétrica de $n \times n$. Dados dos vectores columna X y $Y \in \mathbf{R}^n$, defina

$$\langle X, Y \rangle = {}^tXAY.$$

(a) Muestre que este símbolo satisface las primeras tres propiedades de un producto escalar.

(b) Dé un ejemplo de una matriz de 1×1 y una matriz no nula de 2×2 tales que no se satisfaga la cuarta propiedad. Si esta cuarta propiedad se satisface, esto es, ${}^tXAX > 0$ para todo $X \neq O$, entonces la matriz A se denomina **definitivamente positiva.**

(c) Dé un ejemplo de una matriz de 2×2 que sea simétrica y definitivamente positiva.

(d) Sea $a > 0$, y sea

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}.$$

Pruebe que A es definitivamente positiva si, y sólo si, $ad - b^2 > 0$. [*Sugerencia:* Sea $X = {}^t(x, y)$ y complete el cuadrado en la expresión ${}^tXAX$.]

(e) Si $a < 0$, muestre que A no es definitivamente positiva.

3. Determine si las siguientes matrices son definitivamente positivas.

(a) $\begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} -2 & 1 \\ 1 & 5 \end{pmatrix}$

(c) $\begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$ (d) $\begin{pmatrix} 4 & 4 \\ 4 & 1 \end{pmatrix}$

(e) $\begin{pmatrix} 4 & 1 \\ 1 & 10 \end{pmatrix}$ (f) $\begin{pmatrix} 4 & -1 \\ -1 & 10 \end{pmatrix}$

La traza de una matriz

4. Sea A una matriz de $n \times n$. Defina la traza de A como la suma de los elementos diagonales. Así, si $A = (a_{ij})$, entonces

$$\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Por ejemplo, si

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},$$

entonces $\operatorname{tr}(A) = 1 + 4 = 5$. Si

$$A = \begin{pmatrix} 1 & -1 & 5 \\ 2 & 1 & 3 \\ 1 & -4 & 7 \end{pmatrix},$$

entonces $\operatorname{tr}(A) = 9$. Calcule la traza de las siguientes matrices:

(a) $\begin{pmatrix} 1 & 7 & 3 \\ -1 & 5 & 2 \\ 2 & 3 & -4 \end{pmatrix}$ (b) $\begin{pmatrix} 3 & -2 & 4 \\ 1 & 4 & 1 \\ -7 & -3 & -3 \end{pmatrix}$ (c) $\begin{pmatrix} -2 & 1 & 1 \\ 3 & 4 & 4 \\ -5 & 2 & 6 \end{pmatrix}$.

5. (a) Demuestre que $\operatorname{tr}(A) = \operatorname{tr}({}^tA)$ para cualquier matriz cuadrada A.
   (b) Demuestre que la traza es una aplicación lineal.

6. Si A es una matriz cuadrada simétrica, demuestre que $\operatorname{tr}(AA) \geq 0$.

7. Sean A y B las matrices indicadas. Muestre que

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

(a) $A = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 4 & 1 \\ 3 & 0 & 1 \end{pmatrix}$ $B = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$

(b) $A = \begin{pmatrix} 1 & 7 & 3 \\ -1 & 5 & 2 \\ 2 & 3 & -4 \end{pmatrix}$ $B = \begin{pmatrix} 3 & -2 & 4 \\ 1 & 4 & 1 \\ -7 & -3 & 2 \end{pmatrix}$

8. (a) Pruebe en general que, si A y B son matrices cuadradas de $n \times n$, entonces

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

   (b) Si C es una matriz de $n \times n$ que tiene una inversa, entonces $\operatorname{tr}(C^{-1}AC) = \operatorname{tr}(A)$.

9. Sea V el espacio vectorial de las matrices simétricas de $n \times n$. Para A y $B \in V$, defina el símbolo

$$\langle A, B \rangle = \operatorname{tr}(AB),$$

donde tr es la traza (suma de los elementos diagonales). Demuestre que las propiedades anteriores, en particular, implican que ésta define un producto escalar definitivamente positivo sobre V.

Los Ejercicios 10 a 13 tratan del producto escalar en el contexto referente al cálculo.

10. Sea V el espacio de las funciones continuas sobre $[0, 2\pi]$ y suponga que el producto escalar está dado por la integral sobre este intervalo como aparece en el texto, esto es,

$$\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)\,dx.$$

Sea $g_n(x) = \cos nx$ para $n \geq 0$, y $h_m(x) = \operatorname{sen} mx$ para $m \geq 1$.

(a) Muestre que $\|g_0\| = \sqrt{2\pi}$, $\|g_n\| = \|h_n\| = \sqrt{\pi}$ para $n \geq 1$.

(b) Muestre que $g_n \perp g_m$ si $m \neq n$, y $g_n \perp h_m$ para todos los m, n. *Sugerencia:* Use fómulas como las siguientes:

$$\operatorname{sen} A \cos B = \tfrac{1}{2}[\operatorname{sen}(A+B) + \operatorname{sen}(A-B)]$$

$$\cos A \cos B = \tfrac{1}{2}[\cos(A+B) + \cos(A-B)].$$

11. Sea $f(x) = x$ en el intervalo $[0, 2\pi]$. Encuentre $\langle f, g_n\rangle$ y $\langle f, h_n\rangle$ para las funciones g_n y h_n del Ejercicio 10. Encuentre los coeficientes de Fourier de f con respecto a g_n y h_n.

12. La misma pregunta que en el Ejercicio 11 pero con $f(x) = x^2$. (Los Ejercicios 10 y 11 brindan al lector un repaso de algunas integrales elementales del cálculo.)

13. (a) Sea $f(x) = x$ en el intervalo $[0, 2\pi]$. Encuentre $\|f\|$.
(b) Sea $f(x) = x^2$ en el mismo intervalo. Encuentre $\|f\|$.

VI, §2. Bases ortogonales

A todo lo largo de esta sección consideramos que V es un espacio vectorial con un producto escalar definitivamente positivo. Se dice que una base $\{v_1, \ldots, v_n\}$ de V es **ortogonal** si sus elementos son mutuamente perpendiculares, i.e., si $\langle v_i, v_j\rangle = 0$, siempre que $i \neq j$. Si además cada elemento de la base tiene norma igual a 1, entonces la base se llama **ortonormal**.

Ejemplo 1. Los vectores unitarios canónicos

$$E_1, \ldots, E_n \qquad \text{en} \qquad \mathbf{R}^n$$

forman una base ortonormal de $\mathbf{R}^n$. Ciertamente, cada uno tiene norma igual a 1 y son mutuamente ortogonales, esto es,

$$E_i \cdot E_j = 0 \qquad \text{si} \qquad i \neq j.$$

Por supuesto, hay muchas otras bases ortonormales de $\mathbf{R}^n$.

En el siguiente sentido, este ejemplo es representativo.

Sea $\{e_1, \ldots, e_n\}$ una base ortonormal de V. Cualquier vector $v \in V$ se puede escribir en términos de coordenadas de la manera siguiente:

$$v = x_1e_1 + \cdots + x_ne_n \qquad \text{donde} \qquad x_i \in \mathbf{R}.$$

Denotemos con w otro elemento de V y escribamos

$$w = y_1e_1 + \cdots + y_ne_n \qquad \text{donde} \qquad y_i \in \mathbf{R}.$$

Entonces

$$\begin{aligned}\langle v, w\rangle &= \langle x_1e_1 + \cdots + x_ne_n, y_1e_1 + \cdots + y_ne_n\rangle \\ &= \textstyle\sum_{i,j=1}^{n}\langle x_ie_i, y_je_j\rangle \\ &= \textstyle\sum_{i=1}^{n} x_iy_i\end{aligned}$$

debido a que $\langle e_i, e_j\rangle = 0$ para $i \neq j$. En consecuencia, si X es la n-tupla de coordenadas de v y Y es la n-tupla de coordenadas de w, entonces

$$\langle v, w\rangle = X \cdot Y$$

de manera que el producto escalar está dado precisamente como el producto interior de las coordenadas. Éste es uno de los usos de las bases ortonormales: identificar el producto escalar con el conocido producto interior.

Ejemplo 2. Considere $\mathbf{R}^2$. Sean

$$A = (1, 1) \quad \text{y} \quad B = (1, -1).$$

Entonces $A \cdot B = 0$, de manera que A es ortogonal a B, y A y B son linealmente independientes. Por tanto, forman una base de $\mathbf{R}^2$ y, de hecho, forman una base ortogonal de $\mathbf{R}^2$. Para obtener de ellos una base ortonormal, dividimos a cada uno entre su norma, de manera que una base ortonormal está dada por

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \quad \text{y} \quad \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right).$$

En general, suponga que tenemos un subespacio W de $\mathbf{R}^n$, y sea $A_1, \ldots, A_r$ cualquier base de W. Queremos obtener una base ortogonal de W, para lo cual aplicamos un proceso de ortogonalización paso a paso. Comenzamos con $A_1 = B_1$. Luego tomamos A_2 y restamos su proyección a lo largo de A_1 para obtener un vector B_2. Luego tomamos A_3 y restamos su proyecciones a lo largo de B_1 y B_2 para obtener un vector B_3. Después tomamos A_4 y restamos sus proyecciones a lo largo de B_1, B_2 y B_3 para obtener un vector B_4. Continuamos de esta manera, lo que finalmente nos llevará a una base ortogonal de W.

Enunciamos este hecho como un teorema y lo probamos en el contexto de espacios vectoriales con un producto escalar.

Teorema 2.1. *Sea V un espacio vectorial de dimensión finita, con un producto escalar definitivamente positivo. Sea W un subespacio de V y sea $\{w_1, \ldots, w_m\}$ una base ortogonal de W. Si $W \neq V$, entonces existen elementos $w_{m+1}, \ldots, w_n$ de V tales que $\{w_1, \ldots, w_n\}$ es una base ortogonal de V.*

Demostración. El método usado en la demostración es tan importante como el teorema, y se conoce como **proceso de ortogonalización de Gram-Schmidt.** Sabemos por el Capítulo III, sección §3, que podemos encontrar elementos $v_{m+1}, \ldots, v_n$ de V tales que

$$\{w_1, \ldots, w_m, v_{m+1}, \ldots, v_n\}$$

es una base de V. Por supuesto, no es una base ortogonal. Sea W_{m+1} el espacio generado por $w_1 \dots, w_m, v_{m+1}$. Primero obtendremos una base ortogonal de W_{m+1}. La idea consiste en considerar v_{m+1} y restarle su proyección a lo largo de $w_1, \dots, w_m$. Así, hagamos

$$c_1 = \frac{\langle v_{m+1}, w_1 \rangle}{\langle w_1, w_1 \rangle}, \quad \dots, \quad c_m = \frac{\langle v_{m+1}, w_m \rangle}{\langle w_m, w_m \rangle}.$$

Sea

$$w_{m+1} = v_{m+1} - c_1 w_1 - \dots - c_m w_m.$$

Entonces w_{m+1} es perpendicular a $w_1, \dots, w_m$. Además, $w_{m+1} \neq O$ (en caso contrario v_{m+1} sería linealmente dependiente de $w_1, \dots, w_m$), y v_{m+1} pertenece al espacio generado por $w_1, \dots, w_{m+1}$ debido a que

$$v_{m+1} = w_{m+1} + c_1 w_1 + \dots + c_m w_m.$$

En consecuencia, $\{w_1, \dots, w_{m+1}\}$ es una base ortogonal de W_{m+1}. Podemos proceder ahora por inducción, mostrando que el espacio W_{m+s} generado por

$$w_1, \dots, w_m, v_{m+1}, \dots, v_{m+s}$$

tiene una base ortogonal

$$\{w_1, \dots, w_{m+1}, \dots, w_{m+s}\}$$

donde $s = 1, \dots, n - m$. Esto concluye la prueba.

Corolario 2.2. *Sea V un espacio vectorial de dimensión finita con un producto escalar definitivamente positivo. Suponga que $V \neq \{O\}$. Entonces V tiene una base ortogonal.*

Demostración. Por hipótesis, existe un elemento v_1 de V tal que $v_1 \neq O$. Supongamos que W es el subespacio generado por v_1 y apliquemos el teorema para obtener la base deseada.

Resumamos el procedimiento del Teorema 2.1 una vez más. Suponga que se nos da una base arbitraria $\{v_1, \dots, v_n\}$ de V. Deseamos ortogonalizarla. Procedamos de la manera siguiente. Hagamos

$$v_1' = v_1,$$

$$v_2' = v_2 - \frac{\langle v_2, v_1' \rangle}{\langle v_1', v_1' \rangle} v_1',$$

$$v_3' = v_3 - \frac{\langle v_3, v_2' \rangle}{\langle v_2', v_2' \rangle} v_2' - \frac{\langle v_3, v_1' \rangle}{\langle v_1', v_1' \rangle} v_1',$$

$$\vdots \qquad \vdots$$

$$v_n' = v_n - \frac{\langle v_n, v_{n-1}' \rangle}{\langle v_{n-1}', v_{n-1}' \rangle} v_{n-1}' - \dots - \frac{\langle v_n, v_1' \rangle}{\langle v_1', v_1' \rangle} v_1'.$$

Entonces $\{v_1', \dots, v_n'\}$ es una base ortogonal.

Dada una base ortogonal, siempre podemos obtener una base ortonormal dividiendo cada vector entre su norma.

Ejemplo 3. Encuentre una base ortonormal para el espacio vectorial generado por los vectores $(1,1,0,1)$, $(1,-2,0,0)$ y $(1,0,-1,2)$.

Denotemos estos vectores con A, B y C. Sea

$$B' = B - \frac{B \cdot A}{A \cdot A} A.$$

En otras palabras, restamos de B su proyección a lo largo de A. Entonces B' es perpendicular a A. Encontramos que

$$B' = \tfrac{1}{3}(4,-5,0,1).$$

Ahora restemos de C su proyección a lo largo de A y B', así que hacemos

$$C' = C - \frac{C \cdot A}{A \cdot A} A - \frac{C \cdot B'}{B' \cdot B'} B'.$$

Como A y B' son perpendiculares entre sí, al considerar el producto escalar de C' con A y B' se muestra que C' es perpendicular tanto a A como a B'. Encontramos que

$$C' = \tfrac{1}{7}(-4,-2,-1,6).$$

Los vectores A, B' y C' son no nulos y mutuamente perpendiculares. Pertenecen al espacio generado por A, B y C. En consecuencia, constituyen una base ortogonal para ese espacio. Si deseamos una base ortonormal, dividimos estos vectores entre su norma, y así obtenemos

$$\frac{A}{\|A\|} = \frac{1}{\sqrt{3}}(1,1,0,1),$$

$$\frac{B'}{\|B'\|} = \frac{1}{\sqrt{42}}(4,-5,0,1),$$

$$\frac{C'}{\|C'\|} = \frac{1}{\sqrt{57}}(-4,-2,-1,6),$$

como una base ortonormal.

Ejemplo 4. Encuentre una base ortogonal para el espacio de soluciones de la ecuación lineal

$$3x - 2y + z = 0.$$

Primero encontramos una base, no necesariamente ortogonal. Por ejemplo, demos a z un valor arbitrario, digamos $z = 1$. Por tanto, tenemos que satisfacer

$$3x - 2y = -1.$$

A simple vista, hacemos $x = 1$, $y = 2$ o bien $x = 3$, $y = 5$, esto es,

$$A = (1,2,1) \qquad \text{y} \qquad B = (3,5,1).$$

Entonces se puede verificar fácilmente que A y B son linealmente independientes. Por el Teorema 4.3 del Capítulo 4, el espacio de soluciones tiene dimensión

2, de manera que A y B forman una base de ese espacio de soluciones. Para obtener una base ortogonal, comenzamos con A. Entonces hacemos

$$\begin{aligned} C &= B - \text{proyección de } B \text{ a lo largo de } A \\ &= B - \frac{B \cdot A}{A \cdot A} A \\ &= \left(\frac{2}{3}, \frac{1}{3}, \frac{-4}{3}\right). \end{aligned}$$

Entonces $\{A, C\}$ es una base ortogonal del espacio de soluciones. Algunas veces es conveniente eliminar el denominador. Podemos usar

$$A = (1, 2, 1) \qquad \text{y} \qquad D = (2, 1, -4)$$

en forma satisfactoria como una base ortogonal de ese espacio. Como comprobación, sustituya en la ecuación original para ver que estos vectores dan soluciones, y también verifique que $A \cdot D = 0$, de manera que son perpendiculares entre sí.

Ejemplo 5. Encuentre una base ortogonal para el espacio de soluciones de las ecuaciones homogéneas

$$\begin{aligned} 3x - 2y + z + w &= 0, \\ x + y + 2w &= 0. \end{aligned}$$

Sea W el espacio de soluciones de $\mathbf{R}^4$. Entonces W es el espacio ortogonal a los dos vectores

$$(3, -2, 1, 1) \qquad \text{y} \qquad (1, 1, 0, 2).$$

Éstos, por supuesto, son linealmente independientes. (Por ejemplo, mediante distintos argumentos el lector puede probar fácilmente que la matriz

$$\begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix}$$

tiene rango igual a 2). En consecuencia,

$$\dim W = 4 - 2 = 2.$$

A continuación hallamos una base para el espacio de soluciones. Pongamos $w = 1$ y resolvamos el sistema

$$\begin{aligned} 3x - 2y + z &= -1, \\ x + y &= -2, \end{aligned}$$

mediante eliminación ordinaria. Si ponemos $y = 0$ obtenemos una solución con $x = -2$ y

$$z = -1 - 3x + 2y = 5.$$

Si ponemos $y = 1$, obtenemos una solución con $x = -3$ y

$$z = -1 - 3x + 2y = 10.$$

Así, obtenemos las dos soluciones siguientes:

$$A = (-2, 0, 5, 1) \qquad \text{y} \qquad B = (-3, 1, 10, 1).$$

(Para comprobar, sustituya en el sistema de ecuaciones original para ver que no se ha cometido ningún error de cómputo.) Estas dos soluciones son linealmente independientes, debido a que, por ejemplo, la matriz

$$\begin{pmatrix} -2 & 0 \\ -3 & 1 \end{pmatrix}$$

tiene rango 2. En consecuencia, $\{A, B\}$ es una base para el espacio de soluciones. Para encontrar una base ortogonal, ortogonalizamos B, para obtener

$$B' = B - \frac{B \cdot A}{A \cdot A} A = B - \frac{19}{10} A.$$

También podemos eliminar denominadores y hacemos $C = 10B'$, de manera que

$$\begin{aligned} C &= (-30, 10, 100, 10) - (-38, 0, 95, 19) \\ &= (8, 10, 5, -9) \end{aligned}$$

Entonces $\{A, C\}$ es una base ortogonal para el espacio de soluciones. (Compruebe de nuevo sustituyendo en el sistema de ecuaciones, y también verifique la perpendicularidad viendo directamente que $A \cdot C = 0$.)

También se puede encontrar una base ortogonal sin adivinar soluciones por simple vista o por eliminación desde el comienzo, de la siguiente manera.

Ejemplo 6. Encuentre una base para el espacio de soluciones de la ecuación

$$3x - 2y + z = 0.$$

El espacio de soluciones es el espacio ortogonal al vector $(3, -2, 1)$ y, por tanto, tiene dimensión igual a 2. Por supuesto, hay muchas bases para este espacio. Para encontrar una, primero extendemos $(3, -2, 1) = A$ a una base de $\mathbf{R}^3$. Hacemos esto seleccionando vectores B y C de manera que A, B y C sean linealmente independientes. Por ejemplo, tome

$$B = (0, 1, 0)$$

y

$$C = (0, 0, 1).$$

Entonces A, B y C son linealmente independientes. Para ver esto, procedemos como de costumbre. Si a, b y c son números tales que

$$aA + bB + cC = 0,$$

entonces

$$\begin{aligned} 3a \quad\;\; &= 0, \\ -2a + b &= 0, \\ a + c &= 0. \end{aligned}$$

Esto se resuelve fácilmente para ver que $a = b = c = 0$, de modo que A, B y C son linealmente independientes. Ahora debemos ortogonalizar estos vectores.

Sea

$$B' = B - \frac{\langle B, A\rangle}{\langle A, A\rangle}A = \left(\frac{3}{7}, \frac{5}{7}, \frac{1}{7}\right),$$

$$C' = C - \frac{\langle C, A\rangle}{\langle A, A\rangle}A - \frac{\langle C, B'\rangle}{\langle B', B'\rangle}B'$$

$$= (0,0,1) - \tfrac{1}{14}(3,-2,1) - \tfrac{1}{35}(3,5,1).$$

Entonces $\{B', C'\}$ es una base para el espacio de soluciones de la ecuación dada. Como el lector puede ver, este procedimiento es ligeramente más largo que el de adivinar primero, e incluye una ortogonalización más que el Ejemplo 4.

En el Teorema 2.1 obtuvimos una base ortogonal para V al comenzar con una base ortogonal para un subespacio. Ahora observemos la situación de manera más simétrica.

Teorema 2.3. *Sea V un espacio vectorial de dimensión n, con un producto escalar definitivamente positivo. Sea $\{w_1, \ldots, w_r, u_1, \ldots, u_s\}$ una base ortogonal para V. Sea W el subespacio generado por $w_1, \ldots, w_r$ y sea U el subespacio generado por $u_1, \ldots, u_s$. Entonces $U = W^\perp$, o bien, por simetría, $W = U^\perp$. En consecuencia, para cualquier subespacio W de V tenemos la siguiente relación:*

$$\dim W + \dim W^\perp = \dim V.$$

Demostración. Probaremos que $W^\perp \subset U$ y $U \subset W^\perp$, de manera que $W^\perp = U$.

Primero sea $v \in W^\perp$. Existen números a_i $(i = 1, \ldots, r)$ y b_j $(j = 1, \ldots, s)$ tales que

$$v = \textstyle\sum_{i=1}^r a_i w_i + \sum_{j=1}^s b_j u_j.$$

Como v es perpendicular a todos los elementos de W, tenemos, para cualquier $k = 1, \ldots, r$:

$$0 = v \cdot w_k = \sum a_i w_i \cdot w_k + \sum b_j u_j \cdot w_k$$

$$= a_k w_k \cdot w_k$$

debido a que $w_i \cdot w_k = 0$ si $i \neq k$, y $u_j \cdot w_k = 0$ para todo j. Como $w_k \cdot w_k \neq 0$, se infiere que $a_k = 0$ para todo $k = 1, \ldots, r$, de manera que v es una combinación lineal de $u_1, \ldots, u_s$ y $v \in U$. Así, $W^\perp \subset U$.

Recíprocamente, sea $v \in U$, de manera que v es una combinación lineal de $u_1, \ldots, u_s$. Como $\{w_1, \ldots, w_r, u_1, \ldots, u_s\}$ es una base ortogonal de V, se infiere que cada u_j es perpendicular a W, por lo que el propio v es perpendicular a W, de modo que $U \subset W^\perp$. Por consiguiente, hemos probado que $U = W^\perp$.

Por último, el Teorema 2.1 muestra que la situación anterior se aplica a cualquier subespacio W de V y, por la definición de dimensión,

$$\dim V = r + s = \dim W + \dim W^{\perp},$$

concluyendo así la prueba del teorema.

Ejemplo 7. Considere $\mathbf{R}^3$. Sean A y B dos vectores linealmente independientes en $\mathbf{R}^3$. Entonces el espacio de vectores que son perpendiculares tanto a A como a B es un espacio de dimensión 1. Si $\{N\}$ es una base para este espacio, cualquier otra base para este espacio es del tipo $\{tN\}$, donde t es un número $\neq 0$.

De nuevo en $\mathbf{R}^3$, sea N un vector no nulo. El espacio de vectores perpendiculares a N es un espacio de dimensión 2, es decir, un plano que pasa por el origen O.

Observación. El Teorema 2.3 brinda una nueva prueba del hecho de que el rango por renglones de una matriz es igual a su rango por columnas. En efecto, sea $A = (a_{ij})$ una matriz de $m \times n$. Sea S el espacio de soluciones de la ecuación $AX = O$, de manera que $S = \operatorname{Ker} L_A$. Por el Teorema 3.2 del Capítulo IV, tenemos que

$$\dim S + \text{rango por columnas} = n,$$

dado que la imagen de L_A es el espacio generado por las columnas de A.

Por otro lado, S es el espacio de vectores en $\mathbf{R}^n$ perpendiculares a los renglones de A, de manera que, si W es el espacio renglón, entonces $S = W^{\perp}$. Por consiguiente, por el Teorema 2.3 obtenemos

$$\dim S + \text{rango por renglones} = n.$$

Esto prueba que rango por renglones = rango por columnas. De alguna manera, ésta es una prueba conceptual más satisfactoria de la relación que la que se hizo con anterioridad empleando operaciones por renglones y por columnas.

Para concluir esta sección señalaremos cierta notación útil. Sean X y $Y \in \mathbf{R}^n$ y consideremos X y Y como vectores columna. Denotemos con $\langle\ ,\ \rangle$ el producto escalar canónico sobre $\mathbf{R}^n$. Así, por definición,

$$\langle X, Y\rangle = {}^tXY.$$

En forma análoga, sea A una matriz de $n \times n$. Entonces

$$\langle X, AY\rangle = {}^tXAY = {}^t({}^tAX)Y = \langle {}^tAX, Y\rangle.$$

Por tanto, obtenemos la fórmula

$$\boxed{\langle X, AY\rangle = \langle {}^tAX, Y\rangle.}$$

La transpuesta de la matriz A corresponde a transponer A a ${}^tA$ de un lado del producto escalar al otro. Esta notación se usa con frecuencia en aplicaciones, lo cual es una de las razones de que se mencione en esta parte.

Ejercicios VI, §2

1. Encuentre bases ortonormales para los subespacios de $\mathbf{R}^3$ generados por los siguientes vectores:
   (a) $(1, 1, -1)$ y $(1, 0, 1)$,
   (b) $(2, 1, 1)$ y $(1, 3, -1)$.
2. Encuentre una base ortonormal para el subespacio de $\mathbf{R}^4$ generado por los vectores $(1, 2, 1, 0)$ y $(1, 2, 3, 1)$.
3. Encuentre una base ortonormal para el subespacio de $\mathbf{R}^4$ generado por $(1, 1, 0, 0)$, $(1, -1, 1, 1)$ y $(-1, 0, 2, 1)$.
4. Encuentre una base ortogonal para el espacio de soluciones de las siguientes ecuaciones.
   (a) $\begin{aligned} 2x + y - z &= 0 \\ y + z &= 0 \end{aligned}$ (b) $x - y + z = 0$
   (c) $\begin{aligned} 4x + 7y - \pi z &= 0 \\ 2x - y + z &= 0 \end{aligned}$ (d) $\begin{aligned} x + y + z &= 0 \\ x - y &= 0 \\ y + z &= 0 \end{aligned}$

En los siguientes ejercicios, consideremos el espacio vectorial de las funciones continuas sobre el intervalo $[0, 1]$. Definamos el producto de dos de dichas funciones, f y g, mediante la regla siguiente:

$$\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt.$$

5. Sea V el subespacio de funciones generado por las dos funciones $f(t) = t$ y $g(t) = t^2$. Halle una base ortonormal para V.
6. Sea V el subespacio generado por las tres funciones 1, t, t^2 (donde 1 es la función constante). Halle una base ortonormal para V.
7. Sea V un espacio vectorial de dimensión finita con un producto escalar definitivamente positivo. Sea W un subespacio. Demuestre que
   $$V = W + W^{\perp} \quad \text{y} \quad W \cap W^{\perp} = \{0\}.$$
   En la terminología del capítulo anterior, esto significa que V es la suma directa de W y su complemento ortogonal. [Use el Teorema 2.3.]
8. En el ejercicio 7, demuestre que $(W^{\perp})^{\perp} = W$. ¿Por qué resulta esto inmediato a partir del Teorema 2.3?
9. (a) Sea V el espacio de las matrices simétricas de $n \times n$. Con respecto a A y $B \in V$, defina
   $$\langle A, B\rangle = \operatorname{tr}(AB),$$
   donde tr es la traza (suma de los elementos de la diagonal). Demuestre que esto satisface todas las propiedades de un producto escalar definitivamente positivo. (El lector quizás haya hecho esto como ejercicio en una sección anterior.)
   (b) Sea W el subespacio de las matrices A tales que $\operatorname{tr}(A) = 0$. ¿Cuál es la dimensión del complemento ortogonal de W, relativo al producto escalar que aparece en el inciso (a)? Proporcione una base explícita para este complemento ortogonal.

10. Sea A una matriz simétrica de $n \times n$. Sean X y $Y \in \mathbf{R}^n$ valores propios para A, esto es, suponga que existen números a y b tales que $AX = aX$ y $AY = bY$. Suponga que $a \neq b$, y pruebe que X y Y son perpendiculares entre sí.

VI, §3. Aplicaciones bilineales y matrices

Sean U, V y W espacios vectoriales, y sea

$$\varphi: U \times V \to W$$

una aplicación. Decimos que φ es **bilineal** si, para cada $u \in U$ fija, la aplicación

$$v \mapsto \varphi(u, v)$$

es lineal y, para cada $v \in V$ fija, la aplicación

$$u \mapsto \varphi(u, v)$$

es lineal. La primera condición desarrollada es la siguiente:

$$\varphi(u, v_1 + v_2) = \varphi(u, v_1) + \varphi(u, v_2),$$
$$\varphi(u, cv) = c\varphi(u, v),$$

y del mismo modo para la segunda condición en el otro lado.

Ejemplo. Sea A una matriz de $m \times n$, $A = (a_{ij})$. Podemos definir una aplicación

$$\varphi_A: \mathbf{R}^m \times \mathbf{R}^n \to \mathbf{R}$$

al hacer

$$\varphi_A(X, Y) = {}^tXAY,$$

la cual, desarrollada, tiene el siguiente aspecto:

$$(x_1, \ldots, x_m) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Se supone que nuestros vectores X y Y son vectores columna, de manera que ${}^tX$ es un vector renglón, tal como se muestra. Entonces ${}^tXA$ es un vector renglón y ${}^tXAY$ es una matriz de 1×1, i.e., un número. Así, φ_A aplica parejas de vectores en los reales. Dicha aplicación φ_A satisface propiedades similares a las de un producto escalar. Si fijamos X, entonces la aplicación $Y \mapsto {}^tXAY$ es lineal, y si fijamos Y, entonces la aplicación $X \mapsto {}^tXAY$ también es lineal. En otras palabras, al fijar X tenemos

$$\varphi_A(X, Y + Y') = \varphi_A(X, Y) + \varphi_A(X, Y'),$$
$$\varphi_A(X, cY) = c\varphi_A(X, Y),$$

y en forma análoga si fijamos Y. Esto sólo es una reformulación de las propiedades de la multiplicación de matrices, a saber,

$$ {}^tXA(Y+Y') = {}^tXAY + {}^tXAY', $$
$$ {}^tXA(cY) = c\,{}^tXAY. $$

En conveniente desarrollar la multiplicación ${}^tXAY$ como suma. Observe que

$$ j\text{-ésima componente de } {}^tXA = \sum_{i=1}^{m} x_i a_{ij}, $$

y, así,

$$ {}^tXAY = \sum_{j=1}^{n}\sum_{i=1}^{m} x_i a_{ij} y_j = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij} x_i y_j. $$

Ejemplo. Sea

$$ A = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix} $$

Si $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ y $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ entonces

$$ {}^tXAY = x_1y_1 + 2x_1y_2 + 3x_2y_1 - x_2y_2. $$

Teorema 3.1. *Dada una aplicación bilineal* $\varphi\colon \mathbf{R}^m \times \mathbf{R}^n \to \mathbf{R}$, *existe una matriz única* A *tal que* $\varphi = \varphi_A$, *i.e., tal que*

$$ \varphi(X,Y) = {}^tXAY. $$

Demostración. El enunciado del Teorema 3.1 es similar al enunciado para representar aplicaciones mediante matrices, y su prueba es una extensión de las pruebas anteriores. Recuerde que usamos las bases canónicas para $\mathbf{R}^n$ a fin de probar estos resultados anteriores, y que usamos coordenadas. En este caso hacemos lo mismo. Sean $E^1,\dots,E^m$ los vectores unitarios canónicos para $\mathbf{R}^m$, y sean $U^1,\dots,U^n$ los vectores unitarios canónicos para $\mathbf{R}^n$. Entonces podemos expresar cualquier $X \in \mathbf{R}^m$ como

$$ X = \sum_{i=1}^{m} x_i E^i $$

y cualquier $Y \in \mathbf{R}^n$ como

$$ Y = \sum_{j=1}^{n} y_j U^j. $$

Entonces,

$$ \varphi(X,Y) = \varphi(x_1E^1 + \cdots + x_mE^m, y_1U^1 + \cdots + y_nU^n). $$

Al usar la linealidad por la izquierda, encontramos que

$$ \varphi(X,Y) = \sum_{i=1}^{m} x_i \varphi(E^i, y_1U^1 + \cdots + y_nU^n). $$

Al usar la linealidad por la derecha, encontramos que

$$ \varphi(X,Y) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_i y_j \varphi(E^i, U^j). $$

Sea

$$ a_{ij} = \varphi(E^i, U^j). $$

Entonces vemos que

$$\varphi(X,Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_i y_j,$$

que es precisamente la expresión que obtuvimos para el producto

$${}^tXAY,$$

donde A es la matriz (a_{ij}). Esto prueba que $\varphi = \varphi_A$ para la elección de a_{ij} dada anteriormente.

También es fácil probar la unicidad, que se puede formular de la manera siguiente.

Unicidad. *Si A y B son matrices de $m \times n$ tales que, para todos los vectores X y Y (de la dimensión apropiada) tenemos*

$${}^tXAY = {}^tXBY,$$

entonces $A = B$.

Demostración. Como la relación anterior es válida para todos los vectores X y Y, es válida en particular para los vectores unitarios, así que apliquemos la relación cuando $X = E^i$ y $Y = U^j$. Entonces la regla para la multiplicación de matrices muestra que

$${}^tE^iAU^j = a_{ij} \qquad \text{y} \qquad {}^tE^iBU^j = b_{ij}.$$

En consecuencia, $a_{ij} = b_{ij}$ para todos los índices i, j. Esto demuestra que $A = B$.

Observación. Las aplicaciones bilineales se pueden sumar y multiplicar por escalares. La suma de dos aplicaciones bilineales es, de nuevo, bilineal, y el producto por un escalar también es bilineal. Por tanto, las aplicaciones bilineales forman un espacio vectorial. Verifique las reglas siguientes:

$$\varphi_{A+B} = \varphi_A + \varphi_B \qquad \text{y} \qquad \varphi_{cA} = c\varphi_A.$$

Entonces el Teorema 3.1 se puede expresar diciendo que la asociación

$$A \mapsto \varphi_A$$

es un isomorfismo entre el espacio de las matrices de $m \times n$ y el espacio de las aplicaciones bilineales de $\mathbf{R}^m \times \mathbf{R}^n$ en $\mathbf{R}$.

Aplicación al cálculo. Si el lector ha estudiado cálculo de varias variables, entonces ha asociado con una función f de n variables la matriz de las segundas derivadas parciales siguientes:

$$\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right).$$

Se puede considerar esta matriz como la matriz asociada con una aplicación bilineal conocida como **hessiano.** Observe que esta matriz es simétrica, puesto

que se puede probar que, para funciones suficientemente suaves, las parciales conmutan, esto es

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Ejercicios VI, §3

1. Sea A una matriz de $n \times n$ y suponga que es simétrica, e.i., $A = {}^tA$. Sea $\varphi_A : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ su aplicación bilineal asociada. Demuestre que

$$\varphi_A(X,Y) = \varphi_A(Y,X)$$

para todos los X y $Y \in \mathbf{R}^n$ y, por tanto, que φ_A es un producto escalar, esto es, que satisface las condiciones **PE 1**, **PE 2** y **PE 3**.

2. Recíprocamente, suponga que A es una matriz de $n \times n$ tal que

$$\varphi_A(X,Y) = \varphi_A(Y,X)$$

para todos los X y Y. Demuestre que A es simétrica.

3. Desarrolle completamente en términos de coordenadas la expresión para ${}^tXAY$ cuando A es la siguiente matriz y X y Y son vectores de la correspondiente dimensión.

(a) $\begin{pmatrix} 2 & -3 \\ 4 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 4 & 1 \\ -2 & 5 \end{pmatrix}$ (c) $\begin{pmatrix} -5 & 2 \\ \pi & 7 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 2 & -1 \\ -3 & 1 & 4 \\ 2 & 5 & -1 \end{pmatrix}$ (e) $\begin{pmatrix} -4 & 2 & 1 \\ 3 & 1 & 1 \\ 2 & 5 & 7 \end{pmatrix}$ (f) $\begin{pmatrix} -\frac{1}{2} & 2 & -5 \\ 1 & \frac{2}{3} & 4 \\ -1 & 0 & 3 \end{pmatrix}$

CAPÍTULO VII

Determinantes

Durante algún tiempo hemos trabajado con vectores y a menudo hemos sentido la necesidad de contar con un método para determinar cuándo los vectores son linealmente independientes. Hasta ahora, el único método disponible para nosotros consistía en resolver un sistema de ecuaciones lineales mediante el método de eliminación. En este capítulo mostraremos un método computacional muy eficiente para resolver ecuaciones lineales y para determinar cuándo los vectores son linealmente independientes.

Los casos correspondientes a los determinantes de 2×2 y 3×3 se desarrollarán por separado y con todo detalle, debido a que el caso general de determinantes de $n \times n$ implica una notación que hace aun más difícil entender los determinantes. Se omitirán algunas pruebas en el caso de $n \times n$.

VII, §1. Determinantes de orden 2

Antes de establecer las propiedades generales de un determinante arbitrario, consideremos un caso especial.

Sea

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

una matriz de 2×2. Definimos su **determinante** como $ad - bc$. Por tanto, el determinante es un número, que denotamos como

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Por ejemplo, el determinante de la matriz

$$\begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$$

es igual a $2 \cdot 4 - 1 \cdot 1 = 7$. El determinante de

$$\begin{pmatrix} -2 & -3 \\ 4 & 5 \end{pmatrix}$$

es igual a $(-2) \cdot 5 - (-3) \cdot 4 = -10 + 12 = 2$.

Se puede considerar el determinante como una función de la matriz A, o también como una función de sus dos columnas. Sean éstas A^1 y A^2, como de costumbre. Entonces escribimos el determinante como

$$D(A), \qquad \operatorname{Det}(A) \qquad \text{o} \qquad D(A^1, A^2).$$

Las siguientes propiedades se verifican con facilidad mediante cálculos directos, que el lector deberá desarrollar por completo.

Propiedad 1. *Como función de los vectores columna, el determinante es lineal.*

Esto significa lo siguiente: suponga por ejemplo que $A^1 = C + C'$ es una suma de dos columnas. Entonces

$$D(C + C', A^2) = D(C, A^2) + D(C', A^2).$$

Si x es un número, entonces

$$D(xA^1, A^2) = xD(A^1, A^2).$$

Es válida una fórmula semejante con respecto a la segunda variable. La fórmula se puede probar en forma directa a partir de la definición de determinante. Por ejemplo, sean b' y d' dos números. Entonces

$$\begin{aligned}
\operatorname{Det}\begin{pmatrix} a & b + b' \\ c & d + d' \end{pmatrix} &= a(d + d') - c(b + b') \\
&= ad + ad' - cb - cb' \\
&= ad - bc + ad' - b'c \\
&= \operatorname{Det}\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \operatorname{Det}\begin{pmatrix} a & b' \\ c & d' \end{pmatrix}.
\end{aligned}$$

Además, si x es un número, entonces

$$\operatorname{Det}\begin{pmatrix} xa & b \\ xc & d \end{pmatrix} = xad - xbc = x(ad - bc) = x \operatorname{Det}\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Empleando la terminología del Capítulo VI, sección §4, podemos decir que *el determinante es bilineal.*

Propiedad 2. *Si las dos columnas son iguales, entonces el determinante es igual a 0.*

Esto es inmediato, puesto que, por hipótesis, el determinante es $ab - ab = 0$.

Propiedad 3. *Si I es la matriz unitaria, $I = (E^1, E^2)$, entonces*

$$D(I) = D(E^1, E^2) = 1.$$

De nuevo, esto es inmediato a partir de la definición $ad - bc$.

Con sólo usar las tres propiedades anteriores podemos probar otras, de la manera siguiente.

Si se suma un múltiplo escalar de una columna a la otra, entonces el valor del determinante no cambia.

En otras palabras, sea x un número. Entonces

$$D(A^1 + xA^2, A^2) = D(A^1, A^2).$$

La prueba es inmediata, a saber:

$$\begin{aligned} D(A^1 + xA^2, A^2) &= D(A^1, A^2) + xD(A^2, A^2) && \text{debido a la linealidad} \\ &= D(A^1, A^2) && \text{debido a la propiedad 2.} \end{aligned}$$

Si se intercambian las dos columnas, entonces el determinante cambia de signo.

En otras palabras, tenemos que $D(A^2, A^1) = -D(A^1, A^2)$ o, desarrollando las componentes,

$$\operatorname{Det}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = -\operatorname{Det}\begin{pmatrix} b & a \\ d & c \end{pmatrix}.$$

Por supuesto que esto se puede probar en forma directa a partir de la fórmula $ad - bc$. Sin embargo, también la deduciremos a partir de la propiedad de que, si las dos columnas son iguales, entonces el determinante es igual a 0. Tenemos:

$$\begin{aligned} 0 &= D(A^1 + A^2, A^1 + A^2) \quad \text{(debido a que cada variable es igual a } A^1 + A^2\text{)} \\ &= D(A^1, A^1 + A^2) + D(A^2, A^1 + A^2) \quad \text{(debido a la linealidad en la primera variable)} \\ &= D(A^1, A^1) + D(A^1, A^2) + D(A^2, A^1) + D(A^2, A^2) \quad \text{(debido a la linealidad en la segunda variable)} \\ &= D(A^1, A^2) + D(A^2, A^1). \end{aligned}$$

Así, vemos que $D(A^2, A^1) = -D(A^1, A^2)$. Observe que esta prueba sólo usa la linealidad en cada variable y el hecho de que

$$D(C, C) = 0$$

si C es un vector.

El determinante de A es igual al determinante de su transpuesta, esto es, $D(A) = D({}^tA)$.

En forma explícita, tenemos que

$$\text{Det}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \text{Det}\begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

Esta fórmula proviene de la fórmula $ad - bc$ para el determinante.

Los vectores A^1 y A^2 son linealmente dependientes si, y sólo si, el determinante es igual a 0.

Daremos una prueba que se ajusta al mismo patrón que aparece en la generalización a espacios de dimensión superior. Primero suponga que A^1 y A^2 son linealmente dependientes, de manera que hay una relación lineal

$$xA^1 + yA^2 = 0$$

donde x y y no son simultáneamente iguales a 0. Digamos que $x \neq 0$, y entonces podemos resolver

$$A^1 = zA^2 \qquad \text{donde} \qquad z = -y/x.$$

Ahora tenemos

$$D(A^1, A^2) = D(zA^2, A^2) = zD(A^2, A^2) = 0$$

mediante el uso de la linealidad y la propiedad de que, si las dos columnas son iguales, el determinante es igual a 0.

Recíprocamente, supongamos que A^1 y A^2 son linealmente independientes. Entonces deben formar una base de $\mathbf{R}^2$, que tiene dimensión 2. En consecuencia, podemos expresar los vectores unitarios E^1 y E^2 como combinaciones lineales de A^1 y A^2, digamos

$$E^1 = xA^1 + yA^2 \qquad \text{y} \qquad E^2 = zA^1 + wA^2,$$

donde x, y, z y w son escalares. Ahora tenemos que

$$\begin{aligned} 1 = D(E^1, E^2) &= D(xA^1 + yA^2, zA^1 + wA^2) \\ &= xzD(A^1, A^1) + xwD(A^1, A^2) + yzD(A^2, A^1) + ywD(A^2, A^2) \\ &= (xw - yz)D(A^1, A^2). \end{aligned}$$

Como este último producto es igual a 1, debemos tener $D(A^1, A^2) \neq 0$. Esto prueba la afirmación deseada.

Por último, probemos la unicidad del determinante mediante un método que funcionará en general:

Teorema 1.1. *Sea φ una función de dos variables vectoriales A^1 y $A^2 \in \mathbf{R}^2$ tal que:*

φ es bilineal, esto es, φ es lineal en cada variable.

$\varphi(A^1, A^1) = 0$ *para todo* $A^1 \in \mathbf{R}^2$.

$\varphi(E^1, E^2) = 1$ *si* E^1 *y* E^2 *son los vectores unitarios canónicos* $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Entonces $\varphi(A^1, A^2)$ *es el determinante.*

Demostración. Escriba

$$A^1 = aE^1 + cE^2 \qquad \text{y} \qquad A^2 = bE^1 + dE^2.$$

Entonces

$$\begin{aligned}
\varphi(A^1, A^2) &= \varphi(aE^1 + cE^2, bE^1 + dE^2) \\
&= ab\varphi(E^1, E^1) + ad\varphi(E^1, E^2) + cb\varphi(E^2, E^1) + cd\varphi(E^2, E^2) \\
&= ad\varphi(E^1, E^2) - bc\varphi(E^1, E^2) \\
&= (ad - bc)\varphi(E^1, E^2) \\
&= ad - bc.
\end{aligned}$$

En cada paso hemos usado una de las propiedades probadas con anterioridad. Esto prueba el teorema.

VII, §2. Determinantes de 3×3 y $n \times n$

Mediante inducción definiremos los determinantes y daremos una fórmula para calcularlos al mismo tiempo. Trabajaremos con el caso de 3×3.

Ya hemos definido los determinantes de 2×2. Sea

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

una matriz de 3×3. Definamos su determinante conforme a la fórmula conocida como **desarrollo por renglones**, digamos el primer renglón. Esto es, definamos

$$(*) \qquad \operatorname{Det}(A) = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

Podemos describir esta suma de la manera siguiente. Sea A_{ij} la matriz obtenida a partir de A al suprimir el renglón i y la columna j. Entonces la suma que expresa a $\operatorname{Det}(A)$ se puede escribir como sigue:

$$a_{11} \operatorname{Det}(A_{11}) - a_{12} \operatorname{Det}(A_{12}) + a_{13} \operatorname{Det}(A_{13}).$$

En otras palabras, cada término consiste en el producto de un elemento del primer renglón por el determinante de la matriz de 2×2 obtenida al suprimir el primer renglón y la columna j y al poner el signo apropiado a este término, tal como se muestra.

Ejemplo 1. Sea

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 4 \\ -3 & 2 & 5 \end{pmatrix}.$$

Entonces

$$A_{11} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \end{pmatrix}, \qquad A_{12} = \begin{pmatrix} 1 & 4 \\ -3 & 5 \end{pmatrix}, \qquad A_{13} = \begin{pmatrix} 1 & 1 \\ -3 & 2 \end{pmatrix}$$

y nuestra fórmula para el determinante de A da por resultado:

$$\begin{aligned} \operatorname{Det}(A) &= 2\begin{vmatrix} 1 & 4 \\ 2 & 5 \end{vmatrix} - 1\begin{vmatrix} 1 & 4 \\ -3 & 5 \end{vmatrix} + 0\begin{vmatrix} 1 & 1 \\ -3 & 2 \end{vmatrix} \\ &= 2(5-8) - 1(5+12) + 0 \\ &= -23. \end{aligned}$$

El determinante de una matriz de 3×3 se puede escribir como

$$D(A) = \operatorname{Det}(A) = D(A^1, A^2, A^3).$$

Usamos esta última expresión si deseamos considerar el determinante como una función de las columnas de A.

Además, no hay ninguna razón particular para seleccionar el desarrollo conforme al primer renglón. También podemos usar el segundo renglón y escribir una suma semejante, a saber:

$$\begin{aligned} &-a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{22}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{23}\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} \\ &\qquad = -a_{21}\operatorname{Det}(A_{21}) + a_{22}\operatorname{Det}(A_{22}) - a_{23}\operatorname{Det}(A_{23}). \end{aligned}$$

De nuevo, cada término es el producto de a_{2j} por el determinante de la matriz de 2×2 obtenido al eliminar el segundo renglón y la columna j y al poner el signo apropiado frente a cada término. Este signo se determina conforme al siguiente patrón:

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}.$$

Se verifica en forma directa que es posible desarrollar el determinante conforme a cualquier renglón al multiplicar todos los términos y desarrollar los determinantes de 2×2, con lo que se obtiene el determinante como una suma alternante de seis términos:

(**)
$$\boxed{\begin{aligned} \operatorname{Det}(A) = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} \\ + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \end{aligned}}$$

También podemos desarrollar conforme a las columnas ajustándonos al mismo principio.

Por ejemplo, el desarrollo conforme a la primera columna:

$$a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

da por resultado precisamente los mismos seis términos que en $(**)$.

En el caso de determinantes de 3×3 tenemos, en consecuencia, el siguiente resultado.

Teorema 2.1. *El determinante satisface la regla para el desarrollo conforme a los renglones y a las columnas, y* $\mathrm{Det}(A) = \mathrm{Det}({}^tA)$. *En otras palabras, el determinante de una matriz es igual al determinante de su transpuesta.*

Ejemplo 2. Calcule el determinante

$$\begin{vmatrix} 3 & 0 & 1 \\ 1 & 2 & 5 \\ -1 & 4 & 2 \end{vmatrix}$$

desarrollando conforme a la segunda columna.

El determinante es igual a

$$2\begin{vmatrix} 3 & 1 \\ -1 & 2 \end{vmatrix} - 4\begin{vmatrix} 3 & 1 \\ 1 & 5 \end{vmatrix} = 2(6-(-1)) - 4(15-1) = -42.$$

Observe que la presencia de un 0 en la segunda columna elimina un término en el desarrollo, puesto que este término sería 0.

También podemos calcular el determinante anterior mediante el desarrollo conforme a la tercera columna, a saber, el determinante es igual a

$$+1\begin{vmatrix} 1 & 2 \\ -1 & 4 \end{vmatrix} - 5\begin{vmatrix} 3 & 0 \\ -1 & 4 \end{vmatrix} + 2\begin{vmatrix} 3 & 0 \\ 1 & 2 \end{vmatrix} = -42.$$

Después, sea $A = (a_{ij})$ una matriz arbitraria de $n \times n$. Sea A_{ij} la matriz de $(n-1) \times (n-1)$ obtenida al eliminar el renglón i y la columna j de A.

$$A_{ij} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{ij} & \cdots & a_{1n} \\ \vdots & \vdots & & | & & \vdots \\ \hline & & & a_{ij} & & \\ \vdots & \vdots & & | & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}$$

Daremos una expresión para el determinante de una matriz de $n \times n$ en términos de determinantes de matrices de $(n-1) \times (n-1)$. Sea i un entero, $1 \leq i \leq n$. Definimos

$$\boxed{D(A) = (-1)^{i+1} a_{i1}\,\mathrm{Det}(A_{i1}) + \cdots + (-1)^{i+n} a_{in}\,\mathrm{Det}(A_{in}).}$$

Esta suma se puede describir con palabras. Para cada elemento del renglón i, tenemos una contribución de un término en la suma. Este término es igual a

$+$ o $-$ el producto de este elemento por el determinante de la matriz obtenida de A al eliminar el rengón i y la columna correspondiente. El signo $+$ o $-$ está determinado conforme al siguiente patrón, tipo tablero de ajedrez:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \end{pmatrix}.$$

Esta suma se conoce como **desarrollo del determinante conforme al i-ésimo renglón.**

Mediante una notación más complicada, que omitimos en este libro, se puede demostrar que el Teorema 2.1 también es válido en el caso de $n \times n$. En particular, el determinante satisface la regla del desarrollo conforme a la j-ésima columna, para cualquier j. Así, tenemos la siguiente fórmula de desarrollo:

$$\boxed{D(A) = (-1)^{i+j} a_{1j} D(A_{1j}) + \cdots + (-1)^{n+j} a_{nj} D(A_{nj}).}$$

En la práctica, para calcular un determinante siempre se emplea un desarrollo conforme a algún renglón o columna.

Para el determinante en el caso de $n \times n$ usamos la misma notación que en los casos de 2×2 o 3×3, a saber,

$$|A| = D(A) = \operatorname{Det}(A) = D(A^1 \ldots, A^n).$$

La notación $D(A^1, \ldots, A^n)$ es especialmente conveniente para denotar el determinante como una función de las columnas, por ejemplo, para establecer el siguiente teorema.

Teorema 2.2. *El determinante satisface las siguientes propiedades:*

1. *Como función de cada vector columna, el determinante es lineal, esto es, si la j-ésima columna A^j es igual a una suma de dos vectores columna, digamos $A^j = C + C'$, entonces*

$$\begin{aligned} &D(A^1, \ldots, C + C', \ldots, A^n) \\ &\qquad = D(A^1, \ldots, C, \ldots, A^n) + D(A^1, \ldots, C', \ldots, A^n). \end{aligned}$$

   Además, si x es un número, entonces

$$D(A^1, \ldots, xA^j, \ldots, A^n) = xD(A^1, \ldots, A^j, \ldots, A^n).$$

2. *Si dos columnas son iguales, esto es, si $A^j = A^k$, donde $j \neq k$, entonces el determinante $D(A)$ es igual a 0.*

3. *Si I es la matriz unitaria, entonces $D(I) = 1$.*

La prueba requiere una notación más complicada y la omitiremos. Se puede desarrollar por inducción y a partir de la fórmula explícita que da el desarrollo del determinante.

Como ejemplo, damos la prueba para el caso de determinantes de 3×3. La prueba es por cálculos directos. Suponga que la primera columna es una suma de dos columnas:

$$A^1 = B + C, \qquad \text{esto es} \qquad \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}.$$

Al sustituir cada término de $(*)$, vemos que cada término se divide en una suma de dos términos que corresponden a B y C. Por ejemplo,

$$a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = b_1 \begin{vmatrix} a_{22} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + c_1 \begin{vmatrix} a_{22} & a_{23} \\ a_{31} & a_{33} \end{vmatrix},$$

$$a_{12} \begin{vmatrix} b_2 + c_2 & a_{23} \\ b_3 + c_3 & a_{33} \end{vmatrix} = a_{12} \begin{vmatrix} b_2 & a_{23} \\ b_3 & a_{33} \end{vmatrix} + a_{12} \begin{vmatrix} c_2 & a_{23} \\ c_3 & a_{33} \end{vmatrix},$$

y del mismo modo para el tercer término. La prueba con respecto a la otra columna es semejante. Además, si x es un número, entonces

$$\begin{aligned} \operatorname{Det}(xA^1, A^2, A^3) &= xa_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} xa_{21} & a_{23} \\ xa_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} xa_{21} & a_{23} \\ xa_{31} & a_{32} \end{vmatrix} \\ &= x \operatorname{Det}(A^1, A^2, A^3) \end{aligned}$$

Después supongamos que dos columnas son iguales, por ejemplo la primera y la segunda, de manera que $A^1 = A^2$. Así,

$$a_{11} = a_{12}, \quad a_{21} = a_{22}, \quad a_{31} = a_{32}.$$

Entonces, de nuevo, el lector puede ver en forma directa que los términos se cancelarán, lo que hace que el determinante sea igual a 0.

Por último, si I es la matriz unitaria de 3×3, entonces $\operatorname{Det}(I) = 1$ tanto si se usa el desarrollo conforme a renglones como si se hace conforme a columnas, debido a que en tal desarrollo todos los términos, excepto uno, son iguales a 0 y este único elemento es igual a 1 multiplicado por el determinante de la matriz unitaria de 2×2, que también es igual a 1.

Una función de varias variables que es lineal en cada variable, esto es, que satisface la primera propiedad de los determinantes, se conoce como **multilineal**. Una función que satisface la segunda propiedad se conoce como **alternante**.

Para calcular determinantes en forma eficiente, necesitamos ciertas propiedades adicionales que serán deducidas sin mayor dificultad a partir de las propiedades **1**, **2** y **3** del Teorema 2.2.

4. *Sean j y k enteros tales que $1 \leq j \leq n$ y $1 \leq k \leq n$ con $j \neq k$. Si se intercambian la columna j y la columna k, entonces el determinante cambia de signo.*

Demostración. En la matriz A reemplacemos la columna j y la columna k por $A^j + A^k$. Obtenemos una matriz con dos columnas iguales, de manera que,

por la propiedad **2**, el determinante es igual a 0. Por la propiedad **1** desarrollamos para obtener:

$$\begin{aligned} 0 &= D(\ldots, A^j + A^k, \ldots, A_j + A^k, \ldots) \\ &= D(\ldots, A^j, \ldots, A^j, \ldots) + D(\ldots, A^j, \ldots, A^k, \ldots) \\ &\quad + D(\ldots, A^k, \ldots, A^j, \ldots) + D(\ldots, A^k, \ldots, A^k, \ldots). \end{aligned}$$

Si usamos la propiedad **2** nuevamente, vemos que dos de estos cuatro términos son iguales a 0 y, en consecuencia, que

$$0 = D(\ldots, A^j \ldots, A^k, \ldots) + D(\ldots, A^k \ldots, A^j, \ldots).$$

En esta última suma, un término debe ser igual a menos el otro, lo que prueba la propiedad 4.

5. *Si se suma un múltiplo escalar de una columna a otra, entonces el valor del determinante no cambia.*

Demostración. Consideremos dos columnas diferentes, digamos las columnas k y j, A^k y A^j con $k \neq j$. Sea x un escalar; sumemos xA^j a A^k. Por la propiedad **1**, el determinante se convierte en

$$D(\ldots, \underset{\overset{\uparrow}{k}}{A^k} + xA^j, \ldots) = D(\ldots, \underset{\overset{\uparrow}{k}}{A^k}, \ldots) + D(\ldots, \underset{\overset{\uparrow}{k}}{A^j}, \ldots)$$

(la letra k señala hacia la columna k). En los dos términos que aparecen a la derecha, la columna indicada se encuentra en el lugar k; pero $D(\ldots, A^k, \ldots)$ es simplemente $D(A)$. Además,

$$D(\ldots, \underset{\overset{\uparrow}{k}}{xA^j}, \ldots) = xD(\ldots, \underset{\overset{\uparrow}{k}}{A^j}, \ldots).$$

Como $k \neq j$, el determinante que aparece a la derecha tiene dos columnas iguales, debido a que A^j se encuentra en el lugar k y también en el lugar j. En consecuencia, es igual a 0. Por tanto,

$$D(\ldots, A^k + xA^j, \ldots) = D(\ldots, A^k, \ldots),$$

con lo que se prueba nuestra propiedad 5.

Puesto que el determinante de una matriz es igual al determinante de su transpuesta, esto es, $\mathrm{Det}(A) = \mathrm{Det}({}^tA)$, obtenemos el siguiente hecho general:

Todas las propiedades establecidas con anterioridad para las operaciones por renglones o por columnas son válidas para **ambas** *operaciones.*

Por ejemplo, si un múltiplo escalar de un renglón se suma a otro renglón, entonces el valor del determinante no cambia.

Con los medios anteriores a nuestra disposición, ahora podemos calcular los determinantes de 3×3 en forma muy eficiente. Al hacerlo aplicamos las operaciones descritas en la propiedad 5. Tratamos de hacer en la matriz A suficientes

componentes iguales a 0. Intentaremos especialmente que todos los elementos, excepto uno, de una columna (o renglón) sean iguales a 0 y luego desarrollaremos conforme a esa columna (o renglón). El desarrollo contendrá sólo un término y nuestro cálculo se reducirá a un determinante de 2×2.

Ejemplo 3. Calcule el siguiente determinante

$$\begin{vmatrix} 3 & 0 & 1 \\ 1 & 2 & 5 \\ -1 & 4 & 2 \end{vmatrix}.$$

Ya tenemos un 0 en el primer renglón. Restemos dos veces el segundo renglón del tercero. Entonces nuestro determinante es igual a

$$\begin{vmatrix} 3 & 0 & 1 \\ 1 & 2 & 5 \\ -3 & 0 & -8 \end{vmatrix}.$$

Desarrollemos conforme a la segunda columna. El desarrollo tan sólo tiene un término $\neq 0$, con un signo $+$, y éste es:

$$2\begin{vmatrix} 3 & 1 \\ -3 & -8 \end{vmatrix}.$$

El determinante de 2×2 se puede evaluar mediante nuestra definición $ad - bc$, y hallamos $2(-24 - (-3)) = -42$.

De igual manera reducimos el cálculo de un determinante de 4×4 al de determinantes de 3×3 y luego al de determinantes de 2×2.

Ejemplo 4. Deseamos calcular el siguiente determinante:

$$\begin{vmatrix} 1 & 3 & 1 & 1 \\ 2 & 1 & 5 & 2 \\ 1 & -1 & 2 & 3 \\ 4 & 1 & -3 & 7 \end{vmatrix}.$$

Sumemos el tercer renglón al segundo y luego sumemos el tercer renglón al cuarto. Esto da por resultado

$$\begin{vmatrix} 1 & 3 & 1 & 1 \\ 3 & 0 & 7 & 5 \\ 1 & -1 & 2 & 3 \\ 4 & 1 & -3 & 7 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 0 & 1 \\ 3 & 0 & 7 & 5 \\ 1 & -1 & 2 & 3 \\ 5 & 0 & -1 & 10 \end{vmatrix}.$$

Después sumemos tres veces el tercer renglón al primero y así obtenemos

$$\begin{vmatrix} 4 & 0 & 7 & 10 \\ 3 & 0 & 7 & 5 \\ 1 & -1 & 2 & 3 \\ 5 & 0 & -1 & 10 \end{vmatrix},$$

lo que desarrollamos conforme a la segunda columna. Sólo hay un término, a saber,

$$\begin{vmatrix} 4 & 7 & 10 \\ 3 & 7 & 5 \\ 5 & -1 & 10 \end{vmatrix};$$

Restemos dos veces el segundo renglón del primero y luego del tercer renglón, lo que da por resultado

$$\begin{vmatrix} -2 & -7 & 0 \\ 3 & 7 & 5 \\ -1 & -15 & 0 \end{vmatrix};$$

lo desarrollamos conforme a la tercera columna y así obtenemos

$$-5(30-7) = -5(23) = -115.$$

Ejercicios VII, §2

1. Calcule los siguientes determinantes.

(a) $\begin{vmatrix} 2 & 1 & 2 \\ 0 & 3 & -1 \\ 4 & 1 & 1 \end{vmatrix}$ (b) $\begin{vmatrix} 3 & -1 & 5 \\ -1 & 2 & 1 \\ -2 & 4 & 3 \end{vmatrix}$ (c) $\begin{vmatrix} 2 & 4 & 3 \\ -1 & 3 & 0 \\ 0 & 2 & 1 \end{vmatrix}$

(d) $\begin{vmatrix} 1 & 2 & -1 \\ 0 & 2 & -1 \\ 0 & 2 & 7 \end{vmatrix}$ (e) $\begin{vmatrix} -1 & 5 & 3 \\ 4 & 0 & 0 \\ 2 & 7 & 8 \end{vmatrix}$ (f) $\begin{vmatrix} 3 & 1 & 2 \\ 4 & 5 & 1 \\ -1 & 2 & -3 \end{vmatrix}$

2. Calcule los siguientes determinantes.

(a) $\begin{vmatrix} 1 & 1 & -2 & 4 \\ 0 & 1 & 1 & 3 \\ 2 & -1 & 1 & 0 \\ 3 & 1 & 2 & 5 \end{vmatrix}$ (b) $\begin{vmatrix} -1 & 1 & 2 & 0 \\ 0 & 3 & 2 & 1 \\ 0 & 4 & 1 & 2 \\ 3 & 1 & 5 & 7 \end{vmatrix}$ (c) $\begin{vmatrix} 3 & 1 & 1 \\ 2 & 5 & 5 \\ 8 & 7 & 7 \end{vmatrix}$

(d) $\begin{vmatrix} 4 & -9 & 2 \\ 4 & -9 & 2 \\ 3 & 1 & 0 \end{vmatrix}$ (e) $\begin{vmatrix} 4 & -1 & 1 \\ 2 & 0 & 0 \\ 1 & 5 & 7 \end{vmatrix}$ (f) $\begin{vmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 8 & 5 & 7 \end{vmatrix}$

(g) $\begin{vmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 27 \end{vmatrix}$ (h) $\begin{vmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 9 \end{vmatrix}$ (i) $\begin{vmatrix} 2 & -1 & 4 \\ 3 & 1 & 5 \\ 1 & 2 & 3 \end{vmatrix}$

3. En general, ¿cuál es el determinante de una matriz diagonal

$$\begin{vmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & & \ddots & 0 \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix}?$$

4. Calcule el determinante $\begin{vmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{vmatrix}$.

5. (a) Sean x_1, x_2, x_3 números. Demuestre que

$$\begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2).$$

*(b) Si $x_1, \ldots, x_n$ son números, entonces muestre por inducción que

$$\begin{vmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ 1 & x_n & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{i<j} (x_j - x_i).$$

El símbolo que aparece a la derecha significa que es el producto de todos los términos $x_j - x_i$, donde $i < j$, e i y j son enteros que varían de 1 a n. Este determinante se conoce como determinante V_n de **Vandermonde**. Para efectuar la inducción en forma sencilla, multiplique cada columna por x_1 y réstela de la siguiente columna a la derecha, comenzando por el lado derecho. El lector encontrará que

$$V_n = (x_n - x_1) \cdots (x_2 - x_1) V_{n-1}.$$

6. Encuentre los determinantes de las siguientes matrices.

(a) $\begin{pmatrix} 1 & 2 & 5 \\ 0 & 1 & 7 \\ 0 & 0 & 3 \end{pmatrix}$ (b) $\begin{pmatrix} -1 & 5 & 20 \\ 0 & 4 & 8 \\ 0 & 0 & 6 \end{pmatrix}$

(c) $\begin{pmatrix} 2 & -6 & 9 \\ 0 & 1 & 4 \\ 0 & 0 & 8 \end{pmatrix}$ (d) $\begin{pmatrix} -7 & 98 & 54 \\ 0 & 2 & 46 \\ 0 & 0 & -1 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & 4 & 6 \\ 0 & 0 & 1 \\ 0 & 0 & 8 \end{pmatrix}$ (f) $\begin{pmatrix} 4 & 0 & 0 \\ -5 & 2 & 0 \\ 79 & 54 & 1 \end{pmatrix}$

(g) $\begin{pmatrix} 1 & 5 & 2 & 3 \\ 0 & 2 & 7 & 6 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix}$ (h) $\begin{pmatrix} -5 & 0 & 0 & 0 \\ 7 & 2 & 0 & 0 \\ -9 & 4 & 1 & 0 \\ 96 & 2 & 3 & 1 \end{pmatrix}$

(i) Sea A una matriz triangular de $n \times n$, digamos una matriz tal que todas las componentes que se encuentran debajo de la diagonal son iguales a 0.

$$A = \begin{pmatrix} a_{11} & & & \\ 0 & a_{22} & & * \\ 0 & 0 & & \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

¿Qué es $D(A)$?

7. Si $a(t), b(t), c(t)$ y $d(t)$ son funciones de t, se puede formar el determinante

$$\begin{vmatrix} a(t) & b(t) \\ c(t) & d(t) \end{vmatrix},$$

igual que con números. Desarrolle por completo el determinante siguiente:

$$\begin{vmatrix} \operatorname{sen} t & \cos t \\ -\cos t & \operatorname{sen} t \end{vmatrix}.$$

8. Desarrolle completamente el siguiente determinante:

$$\begin{vmatrix} t+1 & t-1 \\ t & 2t+5 \end{vmatrix}.$$

9. (*Con cálculo*) Sean $f(t)$ y (t) dos funciones que tienen derivadas de todos los órdenes. Sea $\varphi(t)$ la función que se obtiene al considerar el determinante

$$\varphi(t) = \begin{vmatrix} f(t) & g(t) \\ f'(t) & g'(t) \end{vmatrix}.$$

Demuestre que

$$\varphi'(t) = \begin{vmatrix} f(t) & g(t) \\ f''(t) & g''(t) \end{vmatrix},$$

esto es, la derivada se obtiene al derivar el renglón inferior.

10. (*Con cálculo*) Sea

$$A(t) = \begin{pmatrix} b_1(t) & c_1(t) \\ b_2(t) & c_2(t) \end{pmatrix}$$

una matriz de 2×2 de funciones diferenciables. Sean $B(t)$ y $C(t)$ sus vectores columna. Sea

$$\varphi(t) = \mathrm{Det}(A(t)).$$

Muestre que

$$\varphi'(t) = D(B'(t), C(t)) + D(B(t), C'(t)).$$

11. Sea c un número y sea A una matriz de 3×3. Muestre que

$$D(cA) = c^3 D(A).$$

12. Sea c un número y sea A una matriz de $n \times n$. Muestre que

$$D(cA) = c^n D(A).$$

13. Sean $c_1, \ldots, c_n$ números. ¿En qué difieren los determinantes

$$D(c_1A^1 \ldots, c_nA^n) \qquad \text{y} \qquad D(A^1, \ldots, A^n)?$$

14. Escriba en forma explícita el desarrollo de un determinante de 4×4 conforme al primer renglón y conforme a la primera columna.

VII, §3. El rango de una matriz y subdeterminantes

En esta sección damos un criterio para la independencia lineal mediante el uso de determinantes.

Teorema 3.1. *Sean $A^1, \ldots, A^n$ vectores columna de dimensión n. Son linealmente dependientes si, y sólo si,*

$$D(A^1 \ldots, A^n) = 0.$$

Demostración. Supongamos que $A^1 \ldots, A^n$ son linealmente dependientes, de manera que existe una relación

$$x_1A^1 + \cdots + x_nA^n = O$$

donde los números $x_1, \ldots, x_n$ no son todos nulos. Digamos que $x_j \neq 0$. Al restar y dividir entre x_j podemos encontrar números c_k, con $k \neq j$, tales que

$$A^j = \sum_{k \neq j} c_k A^k.$$

Por tanto,

$$D(A) = D\left(A^1 \ldots, \sum_{k\neq j} c_k A^k \ldots, A^n\right)$$
$$= \sum_{k\neq j} D(A^1 \ldots, A^k \ldots, A^n)$$

donde A^k aparece en el lugar j. Sin embargo, A^k también aparece en el lugar k, y $k \neq j$. En consecuencia, el determinante es igual a 0 debido a la propiedad **2**. Esto concluye la prueba de la primera parte.

Por lo que se refiere al recíproco, recordemos que una matriz es equivalente por renglones a una matriz escalonada. Suponga que $A^1, \ldots, A^n$ son linealmente independientes. Entonces la matriz

$$A = (A^1 \ldots, A^n)$$

es equivalente por renglones a una matriz triangular. En realidad, es equivalente por renglones a una matriz B que se encuentra en forma escalonada:

$$\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{pmatrix}$$

y las operaciones de equivalencia por renglones no cambian la propiedad de que los renglones o las columnas sean linealmente independientes. Por tanto, todos los elementos diagonales $b_{11}, \ldots, b_{nn}$ son $\neq 0$. El determinante de esta matriz es el producto

$$b_{11} \cdots b_{nn} \neq 0$$

debido a la regla del desarrollo conforme a las columnas. Bajo las operaciones de equivalencia por renglones, la propiedad del determinante de ser $\neq 0$ no cambia porque las equivalencias por renglones implican la multiplicación de un renglón por un escalar no nulo, lo cual multiplica al determinante por este escalar; o el intercambio de renglones, lo cual multiplica al determinante por -1; o la suma de un múltiplo de un renglón a otro, lo cual no cambia el valor del determinante. Como $\operatorname{Det}(B) \neq 0$, se infiere que $\operatorname{Det}(A) \neq 0$. Esto concluye la prueba.

Corolario 3.2. *Si $A^1, \ldots, A^n$ son vectores columna de $\mathbf{R}^n$ tales que $D(A^1, \ldots, A^n) \neq 0$ y si B es un vector columna, entonces existen números $x_1, \ldots, x_n$ tales que*

$$x_1 A^1 + \cdots + x_n A^n = B.$$

Estos números están determinados de manera única por B.

Demostración. Conforme al teorema, $A^1, \ldots, A^n$ son linealmente independientes y, en consecuencia, forman una base de $\mathbf{R}^n$. Por tanto, cualquier vector de $\mathbf{R}^n$ se puede escribir como una combinación lineal de $A^1 \ldots, A^n$. Como $A^1 \ldots, A^n$ son linealmente independientes, los números $x_1, \ldots, x_n$ son únicos.

En términos de ecuaciones lineales, este corolario muestra lo siguiente:

Si un sistema de n ecuaciones lineales con n incógnitas tiene una matriz de coeficientes cuyo determinante no es 0, entonces este sistema tiene una solución única.

Puesto que los determinantes se pueden usar para probar la independencia lineal, se pueden usar para determinar el rango de una matriz.

Ejemplo 1. Sea

$$A = \begin{pmatrix} 3 & 1 & 2 & 5 \\ 1 & 2 & -1 & 2 \\ 1 & 1 & 0 & 1 \end{pmatrix}.$$

Ésta es una matriz de 3×4. Su rango es a lo sumo 3. Si podemos hallar tres columnas linealmente independientes, entonces sabremos que su rango es exactamente igual a 3. Pero el determinante

$$\begin{vmatrix} 3 & 1 & 2 \\ 1 & 2 & -1 \\ 1 & 1 & 0 \end{vmatrix}$$

no es igual a 0 (a saber, es igual a -3, como se aprecia al restar la segunda columna de la primera y después desarrollar conforme al último renglón). Por tanto, el rango de $A = 3$.

Puede ser que, en una matriz de 3×4, algún determinante de una submatriz de 3×3 sea igual a 0, aunque la matriz de 3×4 tenga rango 3. Por ejemplo, sea

$$B = \begin{pmatrix} 3 & 1 & 2 & 5 \\ 1 & 2 & -1 & 2 \\ 4 & 3 & 1 & 1 \end{pmatrix}.$$

El determinante de las primeras tres columnas

$$\begin{vmatrix} 3 & 1 & 2 \\ 1 & 2 & -1 \\ 4 & 3 & 1 \end{vmatrix}$$

es igual a 0 (en efecto, el último renglón es la suma de los primeros dos renglones). Sin embargo, el determinante

$$\begin{vmatrix} 1 & 2 & 5 \\ 2 & -1 & 2 \\ 3 & 1 & 1 \end{vmatrix}$$

no es nulo (¿a qué es igual?) de manera que, de nuevo, el rango de B es igual a 3.

Si el rango de una matriz de 3×4

$$C = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \end{pmatrix}$$

es igual a 2 o menos, entonces el determinante de *toda* submatriz de 3×3 debe ser igual a 0; en caso contrario, podríamos proceder como antes para obtener tres columnas linealmente independientes. Observemos que hay cuatro de tales subdeterminantes, obtenidos mediante la eliminación sucesiva de una de las cuatro columnas. Recíprocamente, si todos esos subdeterminantes de toda submatriz de 3×3 son iguales a 0, entonces es fácil ver que el rango es a lo sumo 2; porque, si el rango fuera igual a 3, entonces habría tres columnas linealmente independientes y su determinante no sería igual a 0. Así, podemos calcular dichos subdeterminantes para obtener una estimación del rango y luego usar el método de acierto y error, y un poco de sentido común, para obtener el rango exacto.

Ejemplo 2. Sea

$$C = \begin{pmatrix} 3 & 1 & 2 & 5 \\ 1 & 2 & -1 & 2 \\ 4 & 3 & 1 & 7 \end{pmatrix}.$$

Si calculamos todo subdeterminante de 3×3, hallaremos 0. En consecuencia, el rango de C es, a lo sumo, igual a 2. Sin embargo, los primeros dos renglones son linealmente independientes, debido a que, por ejemplo, el determinante

$$\begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix}$$

no es igual a 0. Es el determinante de las primeras dos columnas de la matriz de 2×4

$$\begin{pmatrix} 3 & 1 & 2 & 5 \\ 1 & 2 & -1 & 2 \end{pmatrix}.$$

En consecuencia, el rango es igual a 2.

Por supuesto, si observamos que el último renglón de C es igual a la suma de los dos primeros, entonces vemos fácilmente que el rango es ≤ 2.

Ejercicios VII, §3

Calcule los rangos de las siguientes matrices.

1. $\begin{pmatrix} 2 & 3 & 5 & 1 \\ 1 & -1 & 2 & 1 \end{pmatrix}$

2. $\begin{pmatrix} 3 & 5 & 1 & 4 \\ 2 & -1 & 1 & 1 \\ 5 & 4 & 2 & 5 \end{pmatrix}$

3. $\begin{pmatrix} 3 & 5 & 1 & 4 \\ 2 & -1 & 1 & 1 \\ 8 & 9 & 3 & 9 \end{pmatrix}$

4. $\begin{pmatrix} 3 & 5 & 1 & 4 \\ 2 & -1 & 1 & 1 \\ 7 & 1 & 2 & 5 \end{pmatrix}$

5. $\begin{pmatrix} -1 & 1 & 6 & 5 \\ 1 & 1 & 2 & 3 \\ -1 & 2 & 5 & 4 \\ 2 & 1 & 0 & 1 \end{pmatrix}$

6. $\begin{pmatrix} 2 & 1 & 6 & 6 \\ 3 & 1 & 1 & -1 \\ 5 & 2 & 7 & 5 \\ -2 & 4 & 3 & 2 \end{pmatrix}$

7. $\begin{pmatrix} 2 & 1 & 6 & 6 \\ 3 & 1 & 1 & -1 \\ 5 & 2 & 7 & 5 \\ 8 & 3 & 8 & 4 \end{pmatrix}$ 8. $\begin{pmatrix} 3 & 1 & 1 & -1 \\ -2 & 4 & 3 & 2 \\ -1 & 9 & 7 & 3 \\ 7 & 4 & 2 & 1 \end{pmatrix}$

9. *(Con cálculo).* Sean $\alpha_1, \dots, \alpha_n$ números distintos entre sí, $\neq 0$. Demuestre que las funciones

$$e^{\alpha_1 t} \dots, e^{\alpha_n t}$$

son linealmente independientes sobre los números. [*Sugerencia*: Suponga que tenemos una relación lineal

$$c_1 e^{\alpha_1 t} + \cdots + c_n e^{\alpha_n t} = 0$$

con constantes c_i, válida para todo t. En caso de que no todos los c_i sean iguales a 0, podemos suponer, sin pérdida de generalidad, que ninguno de ellos es 0. Derive la relación anterior $n-1$ veces. Obtendrá un sistema de ecuaciones lineales. El determinante de sus coeficientes debe ser nulo. (¿Por qué?) Obtenga una contradicción a partir de este hecho.]

VII, §4. Regla de Cramer

Se puede usar las propiedades de los determinantes a fin de probar una bien conocida regla para resolver ecuaciones lineales.

Teorema 4.1. *Sean $A^1 \dots, A^n$ vectores columna tales que*

$$D(A^1 \dots, A^n) \neq 0.$$

Sea B un vector columna; si $x_1, \dots, x_n$ son números tales que

$$x_1 A^1 + \cdots + x_n A^n = B,$$

entonces, para cada $j = 1 \dots, n$, tenemos que

$$x_j = \frac{D(A^1, \dots, B \dots, A^n)}{D(A^1, \dots, A_n)}$$

donde B aparece en la j-ésima columna en lugar de A^j. Expresado en otra forma,

$$x_j = \frac{\begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ a_{21} & \cdots & b_2 & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}}$$

(El numerador se obtiene de A al reemplazar la j-ésima columna A^j por B. El denominador es el determinante de la matriz A.)

El teorema 4.1 nos da una forma explícita para hallar las coordenadas de B con respecto a $A^1, \dots, A^n$. Usando el lenguaje de las ecuaciones lineales, el

Teorema 4.1 nos permite resolver explícitamente, en términos de determinantes, el sistema de n ecuaciones lineales con n incógnitas:

$$\begin{matrix} x_1a_{11} + \cdots + x_na_{1n} = b_1 \\ \vdots \qquad\quad \vdots \qquad \vdots \\ x_1a_{n1} + \cdots + x_na_{nn} = b_n. \end{matrix}$$

Probemos el Teorema 4.1

Si B se escribe como en el enunciado del teorema, y se considera el determinante de la matriz obtenida al reemplazar la j-ésima columna de A por B, entonces

$$D(A^1, \ldots, B \ldots, A^n) = D(A^1, \ldots, x_1A_1^1 + \cdots + x_nA_n^1, \ldots, A^n).$$

Usamos la propiedad **1** y obtenemos una suma:

$$D(A^1, \ldots, x_1A^1, \ldots, A^n) + \cdots + D(A^1, \ldots, x_jA^j, \ldots, A^n) \\ + \cdots + D(A^1, \ldots, x_nA^n, \ldots, A^n),$$

la cual, de nuevo por la propiedad **1**, es igual a

$$x_1D(A^1, \ldots, A^1, \ldots, A^n) + \cdots + x_jD(A^1, \ldots, A^n) \\ + \cdots + x_nD(A^1, \ldots, A^n, \ldots, A^n).$$

En todos los términos de esta suma, excepto el j-ésimo, dos vectores columna son iguales. En consecuencia, todo término, excepto el j-ésimo, es igual a 0, por la propiedad **2**. El término j es igual a

$$x_jD(A^1 \ldots, A^n),$$

y, por consiguiente, es igual al determinante con el que comenzamos, a saber, $D(A^1, \ldots, B, \ldots, A^n)$. Podemos despejar x_j y obtener precisamente la expresión dada en el enunciado del teorema.

La regla del Teorema 4.1 que nos da la solución del sistema de ecuaciones lineales por medio de determinantes se conoce como **regla de Cramer**.

Ejemplo. Resuelva el sistema de ecuaciones lineales:

$$\begin{aligned} 3x + 2y + 4z &= 1, \\ 2x - \ y + \ z &= 0, \\ x + 2y + 3z &= 1. \end{aligned}$$

Tenemos:

$$x = \frac{\begin{vmatrix} 1 & 2 & 4 \\ 0 & -1 & 1 \\ 1 & 2 & 3 \end{vmatrix}}{\begin{vmatrix} 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{vmatrix}}, \qquad y = \frac{\begin{vmatrix} 3 & 1 & 4 \\ 2 & 0 & 1 \\ 1 & 1 & 3 \end{vmatrix}}{\begin{vmatrix} 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{vmatrix}}, \qquad z = \frac{\begin{vmatrix} 3 & 2 & 1 \\ 2 & -1 & 0 \\ 1 & 2 & 1 \end{vmatrix}}{\begin{vmatrix} 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{vmatrix}}.$$

Observe cómo se desplaza la columna

$$B = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

de la primera columna al resolver para x, a la segunda columna al resolver para y, a la tercera columna al resolver para z. El denominador en las tres expresiones es el mismo, a saber, es el determinante de la matriz de coeficientes de las ecuaciones.

Como ya sabemos calcular los determinantes de 3×3, hallamos que

$$x = -\tfrac{1}{5}, \quad y = 0, \quad z = \tfrac{2}{5}.$$

Ejercicios VII, §4

1. Resuelva los siguientes sistemas de ecuaciones lineales.

(a)
$$\begin{aligned} 3x + y - z &= 0 \\ x + y + z &= 0 \\ y - z &= 0 \end{aligned}$$

(b)
$$\begin{aligned} 2x - y + z &= 1 \\ x + 3y - 2z &= 0 \\ 4x - 3y + z &= 2 \end{aligned}$$

(c)
$$\begin{aligned} 4x + y + z + w &= 1 \\ x - y + 2z - 3w &= 0 \\ 2x + y + 3z + 5w &= 0 \\ x + y - z - w &= 2 \end{aligned}$$

(d)
$$\begin{aligned} x + 2y - 3z + 5w &= 0 \\ 2x + y - 4z - w &= 1 \\ x + y + z + w &= 0 \\ -x - y - z + w &= 4 \end{aligned}$$

VII, §5. Inversa de una matriz

Consideremos primero un caso especial. Sea

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

una matriz de 2×2 y suponga que su determinante $ad - bc \neq 0$. Deseamos hallar una inversa para A, esto es, una matriz de 2×2

$$X = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$$

tal que

$$AX = XA = I.$$

Veamos el primer requisito, $AX = I$; escrito en forma explícita se ve así:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Veamos la primera columna de AX. Debemos resolver las ecuaciones

$$ax + bz = 1,$$

$$cx + dz = 0.$$

Éste es un sistema de dos ecuaciones con dos incógnitas, x y z, que sabemos resolver. En forma análoga, observando la segunda columna vemos que hay que resolver un sistema de ecuaciones con y y w como incógnitas, a saber,

$$ay + bw = 0,$$
$$cy + dw = 1.$$

Ejemplo. Sea

$$A = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}.$$

Busquemos una matrix X tal que $AX = I$. Por consiguiente, debemos resolver los siguientes sistemas de ecuaciones lineales

$$\begin{aligned} 2x + \ \ z &= 1, \\ 4x + 3z &= 0, \end{aligned} \qquad \text{y} \qquad \begin{aligned} 2y + \ \ w &= 0, \\ 4y + 3w &= 1. \end{aligned}$$

Mediante el método ordinario para resolver dos ecuaciones con dos incógnitas hallamos

$$x = 1, \quad z = -1, \qquad \text{y} \qquad y = -\tfrac{1}{2}, \qquad w = 1.$$

Así, la matriz

$$X = \begin{pmatrix} 1 & -\frac{1}{2} \\ -1 & 1 \end{pmatrix}$$

es tal que $AX = I$. El lector verificará también mediante multiplicación directa que $XA = I$. Esto resuelve el problema de encontrar la inversa deseada.

En forma análoga, en el caso de 3×3 hallaríamos tres sistemas de ecuaciones lineales, correspondientes a la primera columna, a la segunda columna y a la tercera columna. Se podría resolver cada sistema para obtener la inversa. Ahora daremos el argumento general.

Sea A una matriz de $n \times n$. Si B es una matriz tal que $AB = I$ y $BA = I$ (I = matriz unitaria de $n \times n$), entonces decimos que B es una **inversa** de A y escribimos $B = A^{-1}$. Si existe una inversa de A, entonces es única. En efecto, sea C una inversa de A, entonces $CA = I$. Al multiplicar por B por la derecha, obtenemos $CAB = B$. Pero $CAB = C(AB) = CI = C$. En consecuencia, $C = B$. Un argumento similar funciona para $AC = I$.

Teorema 5.1. *Sea $A = (a_{ij})$ una matriz de $n \times n$ y suponga que $D(A) \neq 0$; entonces A es invertible. Sea E^j el vector unitario de la columna j y sea*

$$b_{ij} = \frac{D(A^1, \ldots, E^j, \ldots, A^n)}{D(A)},$$

donde E^j aparece en el lugar i. Entonces la matriz $B = (b_{ij})$ es una inversa de A.

Demostración. Sea $X = (X_{ij})$ una matriz indeterminada de $n \times n$. Deseamos resolver para las componentes x_{ij}, de manera que satisfagan $AX = I$. A partir

de la definición de producto de matrices, esto significa que, para cada j, debemos resolver

$$E^j = x_{1j}A^1 + \cdots + x_{nj}A^n.$$

Éste es un sistema de ecuaciones lineales que se puede resolver en forma única mediante la regla de Cramer, y así obtenemos

$$x_{ij} = \frac{D(A^1, \ldots, E^j, \ldots, A^n)}{D(A)},$$

que es la fórmula dada en el teorema.

Aún debemos probar que $XA = I$. Observemos que $D({}^tA) \neq 0$. En consecuencia, por lo que ya hemos probado, podemos encontrar una matriz Y tal que ${}^tAY = I$. Tomando transpuestas, obtenemos ${}^tYA = I$. Ahora tenemos

$$I = {}^tY(AX)A = {}^tYA(XA) = XA,$$

con lo que se prueba lo que queremos, a saber, que $X = B$ es una inversa de A.

Podemos escribir con detalle las componentes de la matriz B que aparecen en el Teorema 5.1 de la manera siguiente:

$$b_{ij} = \frac{\begin{vmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{j1} & \cdots & 1 & \cdots & a_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix}}{\operatorname{Det}(A)}.$$

Si desarrollamos el determinante del numerador conforme a la columna i, entonces todos los términos, excepto uno, son iguales a cero y, por tanto, obtenemos el numerador de b_{ij} como subdeterminante de $\operatorname{Det}(A)$. Sea A_{ij} la matriz obtenida de A al eliminar el renglón i y la columna j. Entonces

$$b_{ij} = \frac{(-1)^{i+j} \operatorname{Det}(A_{ji})}{\operatorname{Det}(A)}$$

(¡observe la inversión de los índices!), y así tenemos la fórmula

$$\boxed{A^{-1} = \text{ transpuesta de } \left(\frac{(-1)^{i+j} \operatorname{Det}(A_{ij})}{\operatorname{Det}(A)} \right).}$$

Una matriz cuadrada cuyo determinante es $\neq 0$ o, lo que es igual, que admite una inversa, se llama **no singular**.

Ejemplo. Encuentre la inversa de la matriz

$$A = \begin{pmatrix} 3 & 1 & -2 \\ -1 & 1 & 2 \\ 1 & -2 & 1 \end{pmatrix}.$$

Por la fórmula, tenemos que

$$ {}^{t}A^{-1} = \left(\frac{(-1)^{i+j}\,\mathrm{Det}(A_{ij})}{\mathrm{Det}(A)} \right). $$

Para $i = 1$, $j = 1$, la matriz A_{11} se obtiene mediante la eliminación del primer renglón y la primera columna, esto es,

$$ A_{11} = \begin{pmatrix} -1 & 2 \\ -2 & 1 \end{pmatrix} $$

y $\mathrm{Det}(A_{11}) = 1 - (-4) = 5$.

Para $i = 1, j = 2$, la matriz A_{12} se obtiene mediante la eliminación del primer renglón y la segunda columna, esto es,

$$ A_{12} = \begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix} $$

y $\mathrm{Det}(A_{12}) = -1 - 2 = -3$.

Para $i = 1, j = 3$, la matriz A_{13} se obtiene mediante la eliminación del primer renglón y la tercera columna, esto es,

$$ A_{13} = \begin{pmatrix} -1 & 1 \\ 1 & -2 \end{pmatrix} $$

y $\mathrm{Det}(A_{13}) = 2 - 1 = 1$.

Podemos calcular $\mathrm{Det}(A) = 16$. Entonces el primer renglón de ${}^{t}A^{-1}$ es

$$ \tfrac{1}{16}(5, 3, 1). $$

Por consiguiente, la primera columna de A^{-1} es

$$ \tfrac{1}{16}\begin{pmatrix} 5 \\ 3 \\ 1 \end{pmatrix}. $$

Observe que el signo cambia debido al patrón de los signos $(-1)^{i+j}$.

Dejamos para el lector el cálculo de las otras columnas de A^{-1}.

Enunciaremos el siguiente teorema sin demostración:

Teorema 5.2. *Para cualesquiera matrices A y B de $n \times n$, el determinante del producto es igual al producto de los determinantes, esto es*

$$ \mathrm{Det}(AB) = \mathrm{Det}(A)\,\mathrm{Det}(B). $$

Entonces, como caso especial, hallamos que, para una matriz invertible A,

$$ \boxed{\mathrm{Det}(A^{-1}) = \mathrm{Det}(A)^{-1}.} $$

En efecto, tenemos que

$$ AA^{-1} = I, $$

y al aplicar la regla para un producto, obtenemos

$$ D(A)D(A^{-1}) = 1, $$

lo que prueba la fórmula para la inversa.

Ejercicios VII, §5

1. Emplee determinantes para encontrar las inversas de las matrices que aparecen en el Capítulo II, sección §5.

2. Describa en forma explícita la inversa de una matriz de 2×2

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

suponiendo que $ad - bc \neq 0$.

VII, §6. Interpretación de los determinantes como área y como volumen

Resulta notable ver cómo el determinante se puede interpretar como volumen. Estudiamos primero el caso de dimensión 2, por lo que hablaremos de área, aunque escribiremos Vol para representar el área de una figura de dimensión 2, con el objeto de mantener la terminología que se generaliza para dimensiones superiores.

Considere el paralelogramo generado por los vectores v y w.

Por definición, este paralelogramo es el conjunto de todas las combinaciones lineales

$$t_1 v + t_2 w \qquad \text{donde} \qquad 0 \leq t_i \leq 1.$$

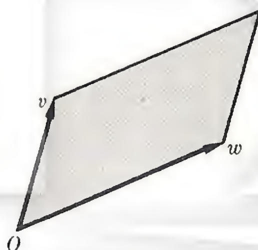


Figura 1

Consideremos v y w como vectores columna y así podemos formar su determinante $D(v, w)$. Este determinante puede ser positivo o negativo, puesto que

$$D(v, w) = D(w, v).$$

Por tanto, el determinante en sí no puede ser el área de este paralelogramo, ya que el área siempre es ≥ 0. Sin embargo, probaremos el siguiente teorema:

Teorema 6.1. *El área del paralelogramo generado por v y w es igual al valor absoluto del determinante, a saber, $|D(v, w)|$.*

Para probar el Teorema 6.1, introduzcamos la noción de área orientada. Sea $P(v, w)$ el paralelogramo generado por v y w. Denotemos con $\mathrm{Vol}_0(v, w)$ el

área de $P(v,w)$ si el determinante $D(v,w) \geq 0$, y menos el área de $P(v,w)$ si el determinante $D(v,w) < 0$. Así, al menos $\text{Vol}_0(v,w)$ tiene el mismo signo que el determinante, y diremos que $\text{Vol}_0(v,w)$ es el **área orientada**. Denotemos con $\text{Vol}(v,w)$ el área del paralegramo generado por v, w. En consecuencia, $\text{Vol}_0(v,w) = \pm\,\text{Vol}(v,w)$.

Para probar el Teorema 6.1, será suficiente probar que:

El área orientada es igual al determinante. Expresado de otro modo,

$$\text{Vol}_0(v,w) = D(v,w).$$

Ahora bien, para probar esto será suficiente demostrar que Vol_0 satisface las tres propiedades características de un determinante, a saber:

1. Vol_0 es lineal en cada variable v y w.
2. $\text{Vol}_0(v,v) = 0$ para todo v.
3. $\text{Vol}_0(E^1, E^2) = 1$ si E^1 y E^2 son los vectores unitarios canónicos.

Sabemos que estas tres propiedades caracterizan a los determinantes, y esto se probó en el Teorema 1.1. A beneficio del lector, repetiremos brevemente el argumento en esta parte. Supongamos que tenemos una función g que satisface estas tres propiedades (g reemplaza a Vol_0). Entonces, para cualesquiera vectores

$$v = aE^1 + cE^2 \qquad \text{y} \qquad w = bE^1 + dE^2$$

tenemos

$$g(aE^1 + cE^2, bE^1 + dE^2) = abg(E^1, E^1) + adg(E^1, E^2) + cbg(E^2, E^1) + cdg(E^2, E^2).$$

El primero y el cuarto términos son iguales a 0. Por el Ejercicio 1,

$$g(E^2, E^1) = -g(E^1, E^2)$$

y, en consecuencia,

$$g(v,w) = (ad - bc)g(E^1, E^2) = ad - bc.$$

Esto prueba lo que queríamos.

Con el objeto de probar que Vol_0 satisface las tres propiedades, usaremos propiedades simples del área (o volumen) como las siguientes: el área de un segmento de recta es igual a 0. Si A es cierta región, entonces el área de A es igual al área de una traslación de A, esto es, igual al área de la región A_w (que consta de todos los puntos $v + w$, con $v \in A$). Si A y B son regiones ajenas entre sí o son tales que sus puntos comunes tienen un área igual a 0, entonces

$$\text{Vol}(A \cup B) = \text{Vol}(A) + \text{Vol}(B).$$

Ahora considere Vol_0. Las últimas dos propiedades son obvias. Ciertamente, el paralelogramo generado por v, v es simplemente un segmento de recta y su área en dos dimensiones es, por consiguiente, igual a 0. Por tanto, se satisface

la propiedad 2. Por lo que se refiere a la tercera propiedad, el paralelogramo generado por los vectores unitarios E^1 y E^2 simplemente es el cuadrado unitario, cuya área es igual a 1. Por ello en este caso tenemos

$$\mathrm{Vol}_0(E^1, E^2) = 1.$$

La propiedad más difícil es la primera. Si es que no lo ha hecho ya, el lector debería leer las aplicaciones geométricas que aparecen en el Capítulo III, sección §2, antes de leer el resto de esta prueba, que se basa en consideraciones geométricas sobre el área.

Necesitaremos el siguiente lema.

Lema 6.2. *Si v y w son linealmente dependientes, entonces* $\mathrm{Vol}_0(v, w) = 0$.

Demostración. Suponga que podemos escribir

$$av + bw = 0$$

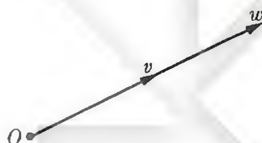


Figura 2

con a o $b \neq 0$; digamos que $a \neq 0$. Entonces

$$v = -\frac{b}{a}w = cw$$

de manera que v y w se encuentran sobre la misma recta y el paralelogramo generado por v, w es un segmento de recta (Fig. 2). En consecuencia, $\mathrm{Vol}_0(v, w) = 0$, lo que prueba el lema.

También sabemos que, cuando v y w son linealmente dependientes, entonces $D(v, w) = 0$, por lo que, en este caso trivial, nuestro teorema está probado. En los lemas subsecuentes supondremos que v y w son linealmente independientes.

Lema 6.3. *Suponga que v y w son linealmente independientes, y sea n un entero positivo. Entonces*

$$\mathrm{Vol}(nv, w) = n\,\mathrm{Vol}(v, w).$$

Demostración. El paralelogramo generado por nv y w consta de n paralelogramos como los que se muestran en la siguiente figura.

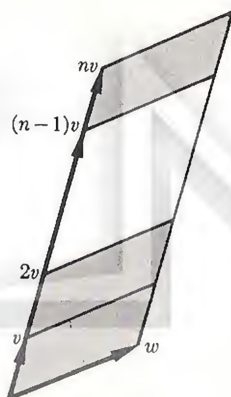


Figura 3

Estos n paralelogramos simplemente son las traslaciones de $P(v, w)$ determinadas por $v, 2v \ldots, (n-1)v$ y cada traslación de $P(v, w)$ tiene la misma área que $P(v, w)$. Estas traslaciones sólo tienen en común segmentos de recta y, por tanto,

$$\mathrm{Vol}(nv, w) = n\,\mathrm{Vol}(v, w)$$

tal como se deseaba.

Corolario 6.4. *Suponga que v y w son linealmente independientes, y sea n un entero positivo. Entonces*

$$\mathrm{Vol}\left(\frac{1}{n}v, w\right) = \frac{1}{n}\,\mathrm{Vol}(v, w).$$

Si m y n son enteros positivos, entonces

$$\mathrm{Vol}\left(\frac{m}{n}v, w\right) = \frac{m}{n}\,\mathrm{Vol}(v, w).$$

Demostración. Sea $v_1 = (1/n)v$. Por el lema sabemos que

$$\mathrm{Vol}(nv_1, w) = n\,\mathrm{Vol}(v_1, w).$$

Esto no es sino una reformulación de nuestro primer aserto, puesto que $nv_1 = v$. Por lo que se refiere a la segunda afirmación, escribimos $m/n = m \cdot 1/n$ y aplicamos en forma sucesiva los enunciados probados:

$$\begin{aligned}\mathrm{Vol}\left(m \cdot \frac{1}{n}v, w\right) &= m\,\mathrm{Vol}\left(\frac{1}{n}v, w\right)\\ &= m \cdot \frac{1}{n}\,\mathrm{Vol}(v, w)\\ &= \frac{m}{n}\,\mathrm{Vol}(v, w).\end{aligned}$$

Lema 6.5. $\text{Vol}(-v, w) = \text{Vol}(v, w)$.

Demostración. El paralelogramo generado por $-v$ y w es una traslación determinada por $-v$ del paralelogramo $P(v, w)$. En consecuencia, $P(v, w)$ y $P(-v, w)$ tiene la misma área (consúltese la Fig. 4).

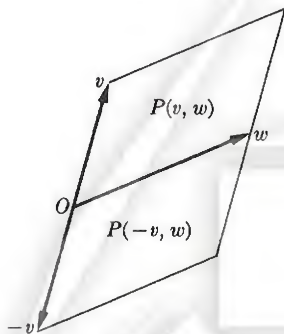


Figura 4

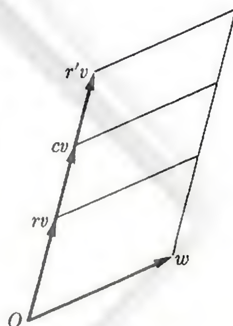


Figura 5

Lema 6.6. *Si c es cualquier número real* > 0, *entonces*

$$\text{Vol}(cv, w) = c\,\text{Vol}(v, w).$$

Demostración. Sean r y r' números racionales tales que $0 < r < c < r'$ (Fig. 5). Entonces

$$P(rv, w) \subset P(cv, w) \subset P(r'v, w).$$

Por tanto, por el Lema 6.3,

$$\begin{aligned} r\,\text{Vol}(v, w) &= \text{Vol}(rv, w) \\ &\leq \text{Vol}(cv, w) \\ &\leq \text{Vol}(r'v, w) \\ &= r'\,\text{Vol}(v, w). \end{aligned}$$

Al hacer que r y r' tiendan a c como límite, encontramos que

$$\text{Vol}(cv, w) = c\,\text{Vol}(v, w),$$

como se quería demostrar.

Ahora, a partir de los lemas 6.5 y 6.6, podemos probar que

$$\boxed{\text{Vol}_0(cv, w) = c\,\text{Vol}_0(v, w)}$$

para cualquier número real c y cualesquiera vectores v y w. Ciertamente, si v y w son linealmente dependientes, entonces ambos lados de la igualdad son iguales a 0. Si v y w son linealmente independientes, usamos la definición de

Vol_0 y los Lemas 6.5 y 6.6; digamos que $D(v,w) > 0$ y c es negativo, $c = -d$. Entonces $D(cv,w) \leq 0$ y, en consecuencia,

$$\begin{aligned}\mathrm{Vol}_0(cv,w) = -\mathrm{Vol}(cv,w) &= -\mathrm{Vol}(-dv,w)\\ &= -\mathrm{Vol}(dv,w)\\ &= -d\,\mathrm{Vol}(v,w)\\ &= c\,\mathrm{Vol}(v,w) = c\,\mathrm{Vol}_0(v,w).\end{aligned}$$

Un argumento similar es válido cuando $D(v,w) \leq 0$. Por consiguiente, hemos probado una de las condiciones de la linealidad de la función Vol_0. Por supuesto, es válida la propiedad análoga para la otra componente, a saber,

$$\boxed{\mathrm{Vol}_0(v,cw) = c\,\mathrm{Vol}_0(v,w).}$$

Para la otra condición, de nuevo tenemos un lema.

Lema 6.7. *Suponga que v y w son linealmente independientes. Entonces*

$$\mathrm{Vol}(v+w,w) = \mathrm{Vol}(v,w).$$

Demostración. Tenemos que probar que el paralelogramo generado por v y w tiene la misma área que el paralelogramo generado por $v+w, w$.

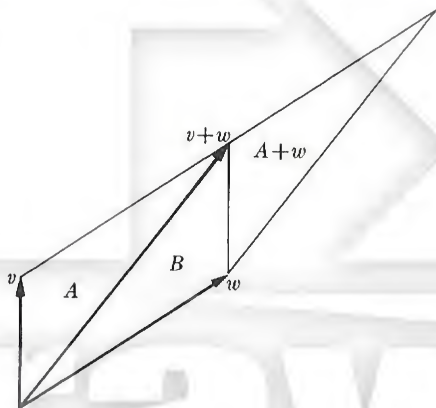


Figura 6

El paralelogramo generado por v y w consta de dos triángulos A y B como se muestra en la figura. El paralelogramo generado por $v+w$ y w consta del triángulo B y la traslación de A determinada por w. Como A y $A+w$ tienen la misma área, obtenemos:

$$\mathrm{Vol}(v,w) = \mathrm{Vol}(A) + \mathrm{Vol}(B) = \mathrm{Vol}(A+w) + \mathrm{Vol}(B) = \mathrm{Vol}(v+w,w),$$

tal como se quería demostrar.

Ahora estamos preparados para trabajar con la segunda propiedad de linealidad. Sea w un vector fijo no nulo en el plano, y sea v un vector tal que $\{v, w\}$ es una base del plano. Probaremos que, para cualesquiera números c y d, tenemos

$$\text{Vol}_0(cv + dw, w) = c\,\text{Vol}_0(v, w). \tag{1}$$

En efecto, si $d = 0$, esto no es más que lo que previamente hemos probado. Si $d \neq 0$, entonces, de nuevo por lo que se demostró previamente,

$$d\,\text{Vol}_0(cv + dw, w) = \text{Vol}_0(cv + dw, dw) = c\,\text{Vol}_0(v, dw) = cd\,\text{Vol}_0(v, w).$$

Al dividir entre d se obtiene la relación (1).

De esta última fórmula se infiere la linealidad. Ciertamente, si

$$v_1 = c_1 v + d_1 w \qquad \text{y} \qquad v_2 = c_2 v + d_2 w,$$

entonces

$$\begin{aligned}
\text{Vol}_0(v_1 + v_2, w) &= \text{Vol}_0\big((c_1 + c_2)v + (d_1 + d_2)w, w\big) \\
&= (c_1 + c_2)\,\text{Vol}_0(v, w) \\
&= c_1\,\text{Vol}_0(v, w) + c_2\,\text{Vol}_0(v, w) \\
&= \text{Vol}_0(v_1, w) + \text{Vol}_0(v_2, w).
\end{aligned}$$

Esto concluye la prueba del hecho de que

$$\text{Vol}_0(v, w) = D(v, w),$$

y, en consecuencia, la del Teorema 6.1.

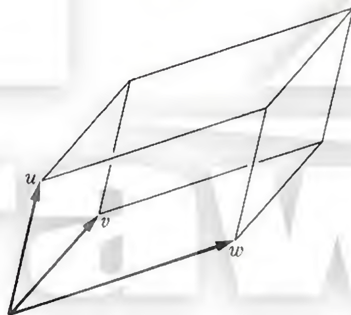


Figura 7

Observación 1. La prueba que se acaba de dar es ligeramente larga, pero todos los pasos son bastante sencillos. Además, cuando se desea generalizar la prueba a un espacio de dimensión superior (incluso al espacio de 3 dimensiones), se puede dar una prueba completamente similar. La razón para ello es que las condiciones que caracterizan a un determinante implican sólo dos coordenadas

a la vez y, así, siempre se lleva a cabo en algún espacio de dimensión 2. Al mantener fijas todas las coordenadas excepto dos, la prueba anterior se puede extender de inmediato. Así, por ejemplo, en el espacio de dimensión 3, denotemos con $P(u, v, w)$ la caja generada por los vectores u, v, w (Fig. 7), a saber, todas las combinaciones

$$t_1u + t_2v + t_3w \qquad \text{donde} \qquad 0 \leq t_i \leq 1.$$

Sea $\text{Vol}(u, v, w)$ el volumen de esta caja.

Teorema 6.8. *El volumen de la caja generada por u, v y w es el valor absoluto del determinante $|D(u, v, w)|$. Esto es,*

$$\text{Vol}(u, v, w) = |D(u, v, w)|.$$

La prueba se ajusta exactamente al mismo patrón que en el caso bidimensional. En efecto, el volumen del cubo generado por los vectores unitarios es igual a 1. Si dos de los vectores u, v y w son iguales, entonces la caja es, en realidad, un paralelogramo de 2 dimensiones cuyo volumen en 3 dimensiones es igual a 0. Por último, la prueba de la linealidad es la misma, debido a que todo el proceso tiene lugar en una o dos variables. Se puede mantener a las otras variables con la misma notación aunque no participen de manera esencial en la prueba.

Asimismo, es posible definir volúmenes de n dimensiones; el teorema correspondiente es como sigue:

Teorema 6.9. *Sean $v_1, \ldots, v_n$ elementos de $\mathbf{R}^n$. Sea $\text{Vol}(v_1, \ldots, v_n)$ el volumen de n dimensiones de la caja de dimensión n generada por $v_1, \ldots, v_n$. Entonces,*

$$\text{Vol}(v_1, \ldots, v_n) = |D(v_1, \ldots, v_n)|.$$

Por supuesto, la caja de n dimensiones generada por $v_1, \ldots, v_n$ es el conjunto de combinaciones lineales

$$\sum_{i=1}^{n} t_i v_i \qquad \text{donde} \qquad 0 \leq t_i \leq 1.$$

Observación 2. Si bien hemos usado propiedades geométricas del área para llevar a cabo la prueba anterior, se pueden sentar los fundamentos en forma puramente analítica para todo esto. Si el lector está interesado, consulte mi libro *Introducción al análisis matemático*, publicado por esta misma editorial.

Observación 3. En el caso especial de 2 dimensiones, en realidad se podría haber dado una prueba más sencilla que la del determinante igual al área, pero preferimos dar una prueba ligeramente más complicada porque es la que se generaliza para el caso de 3 dimensiones o de n dimensiones.

Interpretemos el Teorema 6.1 en términos de aplicaciones lineales. Dados los vectores v y w en el plano, sabemos que existe una única aplicación lineal

$$L : \mathbf{R}^2 \to \mathbf{R}^2$$

tal que $L(E^1) = v$ y $L(E^2) = w$. En efecto, si

$$v = aE^1 + cE^2, \quad w = bE^1 + dE^2,$$

entonces la matriz asociada con la aplicación lineal es

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Además, si denotamos con C el cubo unitario generado por E^1 y E^2 y con P el paralelogramo generado por v y w, entonces P es la imagen bajo L de C, esto es, $L(C) = P$. Ciertamente, tal como hemos visto, para $0 \leq t_i \leq 1$ tenemos

$$L(t_1E^1 + t_2E^2) = t_1L(E^1) + t_2L(E^2) = t_1v + t_2w.$$

Si definimos el determinante de una aplicación lineal como el determinante de su matriz asociada, concluimos que

$$(\text{Área de } P) = |\operatorname{Det}(L)|.$$

Para dar un ejemplo numérico, el área del paralelogramo generado por los vectores (2,1) y (3,−1) (Fig. 8) es igual al valor absoluto de

$$\begin{vmatrix} 2 & 1 \\ 3 & -1 \end{vmatrix} = -5$$

y, por tanto, es igual a 5.

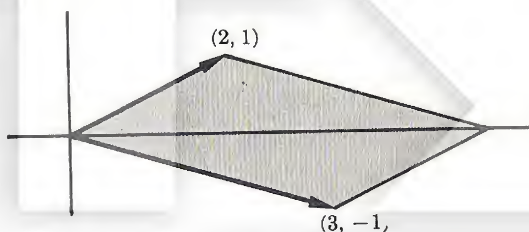


Figura 8

Teorema 6.10. *Sea P un paralelogramo generado por dos vectores. Sea $L : \mathbf{R}^2 \to \mathbf{R}^2$ una aplicación lineal. Entonces*

$$\text{Área de } L(P) = |\operatorname{Det} L|\ (\text{Área de } P).$$

Demostración. Supongamos que P está generado por dos vectores v y w. Entonces $L(P)$ está generado por $L(v)$ y $L(w)$. (Consulte la Fig. 9). Existe una aplicación lineal $L_1 : \mathbf{R}^2 \to \mathbf{R}^2$ tal que

$$L_1(E^1) = v \qquad \text{y} \qquad L_1(E^2) = w.$$

Entonces $P = L_1(C)$, donde C es el cuadrado unitario, y

$$L(P) = L(L_1(C)) = (L \circ L_1)(C).$$

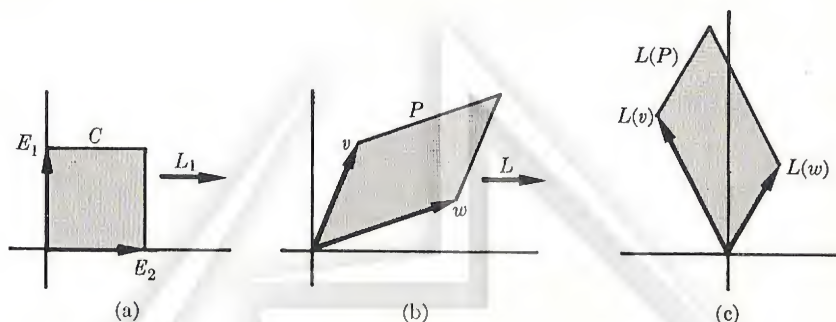


Figura 9

Por lo que hemos probado anteriormente en (*), obtenemos

$$\operatorname{Vol} L(P) = |\operatorname{Det}(L \circ L_1)| = |\operatorname{Det}(L) \operatorname{Det}(L_1)| = |\operatorname{Det}(L)| \operatorname{Vol}(P),$$

lo que prueba nuestro aserto.

Corolario 6.11. *Para cualquier rectángulo R cuyos lados están paralelos a los ejes, y cualquier aplicación lineal $L : \mathbf{R}^2 \to \mathbf{R}^2$, tenemos*

$$\operatorname{Vol} L(R) = |\operatorname{Det}(L)| \operatorname{Vol}(R).$$

Demostración. Sean c_1 y c_2 las longitudes de los lados de R. Sea R_1 el rectángulo generado por $c_1 E^1$ y $c_2 E^2$. Entonces R es la traslación de R_1 determinada por algún vector, digamos $R = R_1 + u$. Entonces

$$L(R) = L(R_1 + u) = L(R_1) + L(u)$$

es la traslación de $L(R_1)$ determinada por $L(u)$. (Consulte la Fig. 10.) Como el área no cambia bajo traslación, sólo necesitamos aplicar el Teorema 6.1 para concluir la prueba.

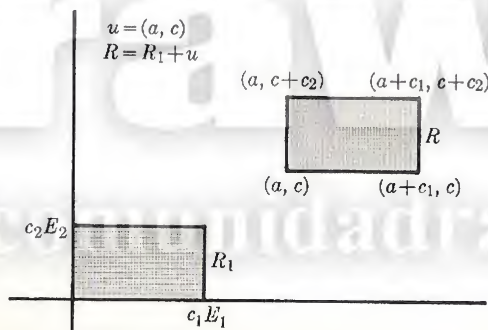


Figura 10

Ejercicios VII, §6

1. Si $g(v, w)$ satisface los dos primeros axiomas de un determinante, pruebe que

$$g(v, w) = -g(w, v)$$

para todos los vectores v y w. Este hecho se usó en la prueba de la unicidad. [*Sugerencia*: Desarrolle $g(v + w, v + w) = 0$.]

2. Encuentre el área del paralelogramo generado por los siguientes vectores.
(a) $(2, 1)$ y $(-4, 5)$ (b) $(3, 4)$ y $(-2, -3)$

3. Encuentre el área del paralelogramo tal que tres de sus vértices se encuentran dados por los siguientes puntos.
(a) $(1, 1), (2, -1), (4, 6)$ (b) $(-3, 2), (1, 4), (-2, -7)$
(c) $(2, 5), (-1, 4), (1, 2)$ (d) $(1, 1), (1, 0), (2, 3)$

4. Encuentre el volumen del paralelepípedo generado por los siguientes vectores en el espacio de 3 dimensiones.
(a) $(1, 1, 3), (1, 2, -1), (1, 4, 1)$ (b) $(1, -1, 4), (1, 1, 0), (-1, 2, 5)$
(c) $(-1, 2, 1), (2, 0, 1), (1, 3, 0)$ (d) $(-2, 2, 1), (0, 1, 0), (-4, 3, 2)$

CAPÍTULO VIII

Vectores propios y valores propios

Este capítulo expone las propiedades elementales básicas de los vectores propios y de los valores propios. Al calcular el polinomio característico se tiene una aplicación de los determinantes. En la sección §3, también se obtiene una interesante combinación del cálculo con el álgebra lineal al relacionar los vectores propios con el problema de encontrar el máximo y el mínimo de una función cuadrática sobre la esfera. La mayoría de los estudiantes que cursan álgebra lineal ha estudiado algo de cálculo, pero si se tiene que evitar el cálculo, se puede usar la prueba que emplea números complejos en lugar del principio del máximo para obtener valores propios reales de una matriz simétrica. En un apéndice se destacan las propiedades básicas de los números complejos.

VIII, §1. Vectores propios y valores propios

Sea V un espacio vectorial y sea

$$A: V \to V$$

una aplicación lineal de V en sí mismo. Se dice que un elemento $v \in V$ es un **vector propio** de A si existe un número λ tal que $Av = \lambda v$. Si $v \neq 0$, *entonces* λ *está determinado en forma única*, debido a que $\lambda_1 v = \lambda_2 v$ implica que $\lambda_1 = \lambda_2$. En este caso, decimos que λ es un **valor propio** de A que pertenece al vector propio v. También decimos que v es un vector propio cuyo valor propio es λ. Además de vector propio y valor propio, también se emplean los términos **vector característico y valor característico.**

Si A es una matriz cuadrada de $n \times n$, entonces un **vector propio** de A es, por definición, un vector propio de la aplicación lineal de $\mathbf{R}^n$ en sí mismo representada por esta matriz. Por tanto, un vector propio X de A es un vector (columna) de $\mathbf{R}^n$ para el cual existe $\lambda \in \mathbf{R}$ tal que $AX = \lambda X$.

Ejemplo 1. Sea V el espacio vectorial sobre $\mathbf{R}$ que consta de todas las funciones infinitamente diferenciables. Sea $\lambda \in \mathbf{R}$; entonces la función f tal que $f(t) = e^{\lambda t}$ es un vector propio de la derivada d/dt, ya que $df/dt = \lambda e^{\lambda t}$.

Ejemplo 2. Sea

$$A = \begin{pmatrix} a_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_n \end{pmatrix}$$

una matriz diagonal. Entonces todo vector unitario E^i $(i = 1, \ldots, n)$ es un vector propio de A. En efecto, tenemos que $AE^i = a_i E$:

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ a_i \\ \vdots \\ 0 \end{pmatrix}.$$

Ejemplo 3. Si $A: V \to V$ es una aplicación lineal y v es un vector propio de A, entonces, para cualquier escalar c no nulo, cv también es un vector propio de A, que tiene el mismo valor propio.

Teorema 1.1. *Sea V un espacio vectorial y sea $A: V \to V$ una aplicación lineal. Sea $\lambda \in \mathbf{R}$ y sea V_λ el subespacio de V generado por todos los vectores propios de A que tienen a λ como valor propio. Entonces todo elemento no nulo de V_λ es un vector propio de A que tiene a λ como valor propio.*

Demostración. Sean v_1 y $v_2 \in V$ tales que $Av_1 = \lambda v_1$ y $Av_2 = \lambda v_2$. Entonces

$$A(v_1 + v_2) = Av_1 + Av_2 = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2).$$

Si $c \in \mathbf{R}$, entonces $A(cv_1) = cAv_1 = c\lambda v_1 = \lambda cv_1$. Esto prueba el teorema.

El subespacio V_λ del Teorema 1.1 se conoce como **espacio propio** de A correspondiente a λ.

Nota. Si v_1 y v_2 son vectores propios de A cuyos valores propios son diferentes, $\lambda_1 \neq \lambda_2$, entonces, por supuesto, $v_1 + v_2$ *no* es un vector propio de A. En efecto, tenemos el siguiente teorema:

Teorema 1.2. *Sea V un espacio vectorial y sea $A: V \to V$ una aplicación lineal. Sean $v_1, \ldots, v_m$ vectores propios de A, cuyos valores propios son $\lambda_1, \ldots, \lambda_m$, respectivamente. Suponga que estos valores propios son distintos entre sí, esto es,*

$$\lambda_i \neq \lambda_j \qquad \text{si} \qquad i \neq j.$$

Entonces $v_1, \ldots, v_m$ son linealmente independientes.

Demostración. Por inducción sobre m. Para $m = 1$, un elemento $v_1 \in V$, $v_1 \neq O$, es linealmente independiente. Suponga que $m > 1$ y que tenemos una relación

$$(*) \qquad c_1 v_1 + \cdots + c_m v_m = O$$

con escalares c_i. Debemos probar que todo $c_i = 0$. Multipliquemos nuestra relación $(*)$ por λ_1 para obtener

$$c_1 \lambda_1 v_1 + \cdots + c_m \lambda_1 v_m = O.$$

También apliquemos A a nuestra relación $(*)$. Por linealidad, obtenemos

$$c_1 \lambda_1 v_1 + \cdots + c_m \lambda_m v_m = O.$$

Ahora restemos estas dos últimas expresiones, con lo que obtenemos

$$c_2(\lambda_2 - \lambda_1)v_2 + \cdots + c_m(\lambda_m - \lambda_1)v_m = O.$$

Como $\lambda_j - \lambda_1 \neq 0$ para $j = 2, \ldots, m$, concluimos por inducción que $c_2 = \cdots = c_m = 0$. Volviendo a nuestra relación original, vemos que $c_1 v_1 = O$, por lo que $c_1 = 0$ y así nuestro teorema quedó probado.

Ejemplo 4. Sea V el espacio vectorial que consta de todas las funciones diferenciables de una variable real t. Sean $\alpha_1, \ldots, \alpha_m$ números distintos entre sí. Las funciones

$$e^{\alpha_1 t}, \ldots, e^{\alpha_m t}$$

son vectores propios de la derivada, con valores propios $\alpha_1, \ldots, \alpha_m$ distintos entre sí y, en consecuencia, son linealmente independientes.

Observación 1. En el Teorema 1.2, suponga que V es un espacio vectorial de dimensión n y que $A: V \to V$ es una aplicación lineal que tiene n vectores propios $v_1, \ldots, v_n$ cuyos valores propios $\lambda_1, \ldots, \lambda_n$ son distintos entre sí. Entonces $\{v_1, \ldots, v_n\}$ es una base de V.

Observación 2. En la teoría de las ecuaciones diferenciales lineales se encuentra una situación como la del Teorema 1.2. Sea $A = (a_{ij})$ una matriz de $n \times n$ y sea

$$F(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}$$

un vector columna de funciones que satisfacen la ecuación

$$\frac{dF}{dt} = AF(t).$$

En términos de coordenadas, esto significa que

$$\frac{df_i}{dt} = \sum_{j=1}^{n} a_{ij} f_j(t).$$

Ahora suponga que A es una matriz diagonal,

$$A = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} \qquad \text{donde } a_i \neq 0 \text{ para todo } i.$$

Entonces toda función $f_i(t)$ satisface la ecuación

$$\frac{df_i}{dt} = a_i f_i(t).$$

Por cálculo sabemos que existen números $c_1, \ldots, c_n$ tales que, para $i = 1, \ldots, n$, tenemos que

$$f_i(t) = c_i e^{a_i t}.$$

[Prueba: si $df/dt = af(t)$, entonces la derivada de $f(t)/e^{at}$ es igual a 0, de manera que $f(t)/e^{at}$ es constante.] Recíprocamente, si $c_1, \ldots, c_n$ son números y hacemos

$$F(t) = \begin{pmatrix} c_1 e^{a_1 t} \\ \vdots \\ c_n e^{a_n t} \end{pmatrix},$$

entonces $F(t)$ satisface la ecuación diferencial

$$\frac{dF}{dt} = AF(t).$$

Sea V el conjunto de soluciones $F(t)$ de la ecuación diferencial $dF/dt = AF(t)$. Entonces se verifica de inmediato que V es un espacio vectorial, y el argumento anterior muestra que los n elementos

$$\begin{pmatrix} e^{a_1 t} \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ e^{a_2 t} \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \begin{pmatrix} 0 \\ 0 \\ \vdots \\ e^{a_n t} \end{pmatrix}$$

forman una base para V. Además, estos elementos son vectores propios de A y también de la derivada (considerada como una aplicación lineal).

Lo anterior es válido si A es una matriz diagonal. Si A no es diagonal, entonces tratamos de encontrar una base tal que podamos representar la aplicación lineal A con una matriz diagonal. Aquí no haremos este tipo de consideraciones.

Ejercicios VIII, §1

Sea a un número $\neq 0$.

1. Pruebe que los vectores propios de la matriz

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$$

generan un espacio de dimensión 1, y dé una base para este espacio.

2. Pruebe que los vectores propios de la matriz

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

generan un espacio de 2 dimensiones y dé una base para este espacio. ¿Cuáles son los valores propios de esta matriz?

3. Sea A una matriz diagonal cuyos elementos diagonales son $a_{11}, \dots, a_{nn}$. ¿Cuál es la dimensión del espacio generado por los vectores propios de A? Muestre una base para este espacio y diga cuáles son los valores propios.

4. Demuestre que, si $\theta \in \mathbf{R}$, entonces la matriz

$$A = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{pmatrix}$$

siempre tiene un vector propio en $\mathbf{R}^2$, y que, de hecho, existe un vector tal que $Av_1 = v_1$. [*Sugerencia*: Tome como primera componente de v_1 la siguiente:

$$x = \frac{\operatorname{sen}\theta}{1 - \cos\theta}$$

si $\cos\theta \neq 1$; luego despeje y. ¿Qué pasa si $\cos\theta = 1$?]

5. En el ejercicio 4, sea v_2 un vector de $\mathbf{R}^2$ perpendicular al vector v_1 hallado en ese ejercicio. Demuestre que $Av_2 = -v_2$ y defina el significado de esto de modo que indique que A es una reflexión.

6. Sea

$$R(\theta) = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

la matriz de una rotación. Demuestre que $R(\theta)$ no tiene valores propios reales, a menos que $R(\theta) = \pm I$. [El lector encontrará más sencillo hacer este ejercicio después de estudiar la siguiente sección.]

7. Sea V un espacio vectorial de dimensión finita. Sean A y B aplicaciones lineales de V en sí mismo. Suponga que $AB = BA$. Demuestre que, si v es un vector propio de A, cuyo valor propio es λ, entonces Bv es un vector propio de A, cuyo valor propio también es λ si $Bv \neq O$.

VIII, §2. El polinomio característico

Ahora veremos cómo se pueden emplear los determinantes para hallar el valor propio de la matriz.

Teorema 2.1. *Sea V un espacio vectorial de dimensión finita y sea λ un número. Sea $A: V \to V$ una aplicación lineal. Entonces λ es un valor propio de A si, y sólo si, $A - \lambda I$ no es invertible.*

Demostración. Supongamos que λ es un valor propio de A. Entonces existe un elemento $v \in V$, $v \neq 0$, tal que $Av = \lambda v$. En consecuencia, $Av - \lambda v = O$, y $(A - \lambda I)v = O$. Por tanto, $A - \lambda I$ tiene un núcleo no nulo, y $A - \lambda I$ no puede ser invertible. Recíprocamente, supongamos que $A - \lambda I$ no es invertible. Por el Teorema 2.4 del Capítulo 5, vemos que $A - \lambda I$ debe tener un núcleo no nulo, lo que significa que existe un elemento $v \in V$, $v \neq O$, tal que $(A - \lambda I)v = O$. En consecuencia, $Av - \lambda v = O$ y $Av = \lambda v$. Así, λ es un valor propio de A. Esto prueba nuestro teorema.

Sea A una matriz de $n \times n$, $A = (a_{ij})$. Definimos el **polinomio característico** P_A de A como el determinante

$$P_A(t) = \mathrm{Det}(tI - A),$$

o, escrito con todo detalle,

$$P(t) = \begin{vmatrix} t - a_{11} & -a_{12} & \cdots & -a_{1n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & t - a_{nn} \end{vmatrix} = \begin{vmatrix} t - a_{11} & & -a_{ij} \\ & \ddots & \\ -a_{ij} & & t - a_{nn} \end{vmatrix}.$$

También podemos considerar A como una aplicación lineal de $\mathbf{R}^n$ en $\mathbf{R}^n$, y decir que $P_A(t)$ es el **polinomio característico de esta aplicación lineal.**

Ejemplo 1. El polinomio característico de la matriz

$$A = \begin{pmatrix} 1 & -1 & 3 \\ -2 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$

es

$$\begin{vmatrix} t-1 & 1 & -3 \\ 2 & t-1 & -1 \\ 0 & -1 & t+1 \end{vmatrix},$$

y lo desarrollamos conforme a la primera columna para obtener

$$P_A(t) = t^3 - t^2 - 4t + 6.$$

Respecto a una matriz arbitraria $A = (a_{ij})$, el polinomio característico se puede hallar desarrollando conforme a la primera columna, y siempre consistirá en una suma

$$(t - a_{11}) \cdots (t - a_{nn}) + \cdots.$$

Todo término distinto del que hemos escrito tendrá grado $< n$. En consecuencia, el polinomio característico es del tipo

$$P_A(t) = t^n + \text{términos de grado menor.}$$

Teorema 2.2. *Sea A una matriz de $n \times n$. Un número λ es un valor propio de A si, y sólo si, λ es una raíz del polinomio característico de A.*

Demostración. Suponga que λ es un valor propio de A. Entonces $\lambda I - A$ no es invertible, debido al Teorema 2.1, y por tanto, $\mathrm{Det}(\lambda I - A) = 0$, por el Teorema 5.1 del Capítulo VII. En consecuencia, λ es una raíz del polinomio característico. Recíprocamente, si λ es una raíz del polinomio característico, entonces

$$\mathrm{Det}(\lambda I - A) = 0,$$

y por tanto, por el mismo Teorema 5.1 del Capítulo VII, concluimos que $\lambda I - A$ no es invertible, de modo que λ es un valor propio de A, según el Teorema 2.1.

El teorema 2.2 nos da una manera explícita de determinar los valores propios de una matriz, *a condición* de que podamos determinar en forma explícita las raíces de su polinomio característico. Algunas veces esto es fácil, en especial en los ejercicios que aparecen al final de los capítulos, donde las matrices se ajustan de tal manera que se puede determinar las raíces a simple vista o mediante recursos sencillos. En otros casos es considerablemente más difícil.

Por ejemplo, para determinar las raíces del polinomio que aparece en el Ejemplo 1, se tendría que desarrollar la teoría de los polinomios cúbicos. Esto se puede hacer, pero implica fórmulas que resultan un tanto más difíciles que la fórmula con la que se resuelve una ecuación cuadrática. También es posible hallar métodos para determinar raíces en forma aproximada. En cualquier caso, la determinación de tales métodos corresponde a un tipo de ideas que no es el que se maneja en este capítulo.

Ejemplo 2. Halle los valores propios y una base para los espacios propios de la siguiente matriz:

$$\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}.$$

El polinomio característico es el determinante

$$\begin{vmatrix} t-1 & -4 \\ -2 & t-3 \end{vmatrix} = (t-1)(t-3) - 8 = t^2 - 4t - 5 = (t-5)(t+1).$$

En consecuencia, los valores propios son 5 y -1.

Para cualquier valor propio λ, un vector propio correspondiente es un vector $\begin{pmatrix} x \\ y \end{pmatrix}$ tal que

$$x + 4y = \lambda x,$$

$$2x + 3y = \lambda y,$$

o, en forma equivalente,

$$(1-\lambda)x + 4y = 0,$$
$$2x + (3-\lambda)y = 0.$$

Demos a x algún valor, digamos $x = 1$, y despejemos y a partir de cualquier ecuación, por ejemplo la segunda, para obtener $y = -2/(3-\lambda)$. Esto nos da el vector propio

$$X(\lambda) = \begin{pmatrix} 1 \\ -2/(3-\lambda) \end{pmatrix}.$$

Al sustituir $\lambda = 5$ y $\lambda = -1$ obtenemos los dos vectores propios

$$X^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{para} \quad \lambda = 5, \qquad \text{y} \qquad X^2 = \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix} \quad \text{para} \quad \lambda = -1.$$

El espacio propio para 5 tiene base X^1 y el espacio propio para -1 tiene base X^2. Observe que cualesquiera múltiplos escalares no nulos de estos vectores también serían bases. Por ejemplo, en lugar de X^2 podríamos considerar

$$\begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Ejemplo 3. Halle los valores propios y una base para los espacios propios de la matriz

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}.$$

El polinomio característico es el determinanate

$$\begin{pmatrix} t-2 & -1 & 0 \\ 0 & t-1 & 1 \\ 0 & -2 & t-4 \end{pmatrix} = (t-2)^2(t-3).$$

Por tanto, los valores propios son 2 y 3.

Para los vectores propios, debemos resolver las ecuaciones siguientes:

$$(2-\lambda)x + y = 0,$$
$$(1-\lambda)y - z = 0,$$
$$2y + (4-\lambda)z = 0.$$

Observe el coeficiente $(2-\lambda)$ de x.

Suponga que queremos encontrar el espacio propio cuyo valor propio es $\lambda = 2$. Entonces la primera ecuación se convierte en $y = 0$, por lo que $z = 0$, a partir de la segunda ecuación. Podemos dar a x cualquier valor, digamos $x = 1$. Entonces el vector

$$X^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

es una base para el espacio propio cuyo valor propio es 2.

Ahora suponga que $\lambda \neq 2$, de manera que $\lambda = 3$. Si ponemos $x = 1$, entonces podemos despejar y a partir de la primera ecuación para obtener $y = 1$, y entonces podemos despejar z en la segunda ecuación para obtener $z = -2$. Por tanto,

$$X^2 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

es una base para los vectores propios cuyo valor propio es 3. Cualquier múltiplo escalar no nulo de X^2 también sería una base.

Ejemplo 4. El polinomio característico de la matriz

$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 5 & -1 \\ 0 & 0 & 7 \end{pmatrix}$$

es $(t-1)(t-5)(t-7)$. ¿Puede el lector generalizar esto?

Ejemplo 5. Halle los valores propios y una base para los espacios propios de la matriz que aparece en el Ejemplo 4.

Los valores propios son 1, 5 y 7. Sea X un vector propio no nulo, digamos

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \qquad \text{que también se escribe} \qquad {}^tX = (x, y, z).$$

Entonces, por definición de vector propio, existe un número λ tal que $AX = \lambda X$, lo que significa que

$$\begin{aligned} x + y + 2z &= \lambda x, \\ 5y - z &= \lambda y, \\ 7z &= \lambda z. \end{aligned}$$

Caso 1. $z = 0$, $y = 0$. Como queremos un vector propio no nulo, debemos tener $x \neq 0$, en cuyo caso $\lambda = 1$, por la primera ecuación. Sea $X^1 = E^1$ el primer vector unitario, o cualquier múltiplo escalar no nulo, para obtener un vector propio cuyo valor propio es 1.

Caso 2. $z = 0$, $y \neq 0$. Por la segunda ecuación, debemos tener $\lambda = 5$. Demos a y un valor específico, digamos $y = 1$. Luego despejemos x de la primera ecuación, a saber,

$$x + 1 = 5x, \qquad \text{lo que da} \qquad x = \tfrac{1}{4}.$$

Sea

$$X^2 = \begin{pmatrix} \frac{1}{4} \\ 1 \\ 0 \end{pmatrix}$$

Entonces X^2 es un vector propio cuyo valor propio es 5.

Caso 3. $z \neq 0$. Entonces, por la tercera ecuación, debemos tener $\lambda = 7$. Fije algún valor no nulo de z, digamos $z = 1$, por lo que todo se reduce a resolver las dos ecuaciones simultáneas

$$x + y + 2 = 7x,$$
$$5y - 1 = 7y.$$

Esto da por resultado $y = -\frac{1}{2}$ y $x = \frac{1}{4}$. Sea

$$X^3 = \begin{pmatrix} \frac{1}{4} \\ -\frac{1}{2} \\ 1 \end{pmatrix}.$$

Entonces X^3 es un vector propio cuyo valor propio es 7.

Los múltiplos escalares de X^1, X^2 y X^3 producirán vectores propios que tienen los mismos valores propios que X^1, X^2 y X^3, respectivamente. Como estos tres vectores tienen distintos valores propios, son linealmente independientes y, por tanto, forman una base de $\mathbf{R}^3$. Por el Ejercicio 14, no hay otros vectores propios.

Por último señalaremos que el álgebra lineal de matrices se podría haber realizado con coeficientes complejos. Lo mismo se puede decir para los determinantes. Todo lo que se necesita de los números es que se puedan sumar, multiplicar y dividir entre números no nulos, y estas operaciones son válidas para los números complejos. Luego, una matriz $A = (a_{ij})$ de números complejos tiene valores propios y vectores propios cuyas componentes son números complejos. Esto es útil por el siguiente hecho fundamental:

Todo polinomio no constante con coeficientes complejos tiene una raíz compleja.

Si A es una matriz compleja de $n \times n$, entonces el polinomio característico de A tiene coeficientes complejos y tiene grado $n \geq 1$, de manera que tiene una raíz compleja que es un valor propio. Por tanto, sobre los números complejos, una matriz cuadrada siempre tiene un valor propio y un vector propio no nulo. Esto no siempre es cierto sobre los números reales. (¿Ejemplo?) En la siguiente sección veremos un importante caso en el que una matriz real siempre tiene un valor propio real.

Ahora daremos algunos ejemplos de cálculos que usan números complejos para los valores propios y vectores propios, aun cuando la propia matriz tenga componentes reales. Recuerde que, en el caso de valores propios complejos, el espacio vectorial está sobre los números complejos, por lo que consta de combinaciones lineales de los elementos de la base dada, con coeficientes complejos.

Ejemplo 6. Halle los valores propios y una base para los espacios propios de la siguiente matriz

$$A = \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix}.$$

El polinomio característico es el determinante

$$\begin{vmatrix} t-2 & 1 \\ -3 & t-1 \end{vmatrix} = (t-2)(t-1)+3 = t^2 - 3t + 5.$$

En consecuencia, los valores propios son

$$\frac{3 \pm \sqrt{9-20}}{2}.$$

Por tanto, hay dos valores propios distintos entre sí (aunque ningún valor propio real):

$$\lambda_1 = \frac{3+\sqrt{-11}}{2} \quad \text{y} \quad \lambda_2 = \frac{3-\sqrt{-11}}{2}.$$

Sea $X = \begin{pmatrix} x \\ y \end{pmatrix}$ donde x y y no son simultáneamente iguales a 0. Entonces X es un vector propio si, y sólo si, $AX = \lambda X$, esto es:

$$2x - y = \lambda x,$$

$$3x + y = \lambda y,$$

donde λ es un valor propio. Este sistema es equivalente a

$$(2-\lambda)x - y = 0,$$

$$3x + (1-\lambda)y = 0.$$

Damos a x, digamos, un valor arbitrario, por ejemplo $x = 1$, y despejamos y, por lo que $y = (2-\lambda)$, por la primera ecuación. Luego obtenemos los vectores propios

$$X(\lambda_1) = \begin{pmatrix} 1 \\ 2-\lambda_1 \end{pmatrix} \quad \text{y} \quad X(\lambda_2) = \begin{pmatrix} 1 \\ 2-\lambda_2 \end{pmatrix}.$$

Observación. A partir de una de las ecuaciones despejamos y. Esto es consistente con la otra ecuación, pues λ es un valor propio. En efecto, si el lector sustituye $x = 1$ y $y = 2-\lambda$ en el miembro de la izquierda de la segunda ecuación, obtendrá

$$3 + (1-\lambda)(2-\lambda) = 0$$

debido a que λ es una raíz del polinomio característico.

Entonces $X(\lambda_1)$ es una base para el espacio propio unidimensional de λ_1 y $X(\lambda_2)$ es una base para el espacio propio unidimensional de λ_2.

Ejemplo 7. Encuentre los valores propios y una base para los espacios propios de la siguiente matriz

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Calculemos el polinomio característico, que es el determinante

$$\begin{vmatrix} t-1 & -1 & 1 \\ 0 & t-1 & 0 \\ -1 & 0 & t-1 \end{vmatrix}$$

el que tras sencillos cálculos da

$$P(t) = (t-1)(t^2 - 2t + 2).$$

Ahora tenemos el problema de hallar las raíces de $P(t)$ como números reales o números complejos. Por la fórmula cuadrática, las raíces de $t^2 - 2t + 2$ están dadas por

$$\frac{2 \pm \sqrt{4-8}}{2} = 1 \pm \sqrt{-1}.$$

Toda la teoría del álgebra lineal se podría desarrollar sobre los números complejos, y los valores propios de la matriz dada también se pueden definir sobre los complejos. Luego, a partir del cálculo de las raíces anteriores, se aprecia que el único valor propio real es 1, y que hay dos valores propios complejos, a saber:

$$1 + \sqrt{-1} \quad \text{y} \quad 1 - \sqrt{-1}.$$

Hacemos que los valores propios sean los siguientes:

$$\lambda_1 = 1, \qquad \lambda_2 = 1 + \sqrt{-1}, \qquad \lambda_3 = 1 - \sqrt{-1}.$$

Sea

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

un vector no nulo. Entonces X es un vector propio para A si, y sólo si, se satisfacen las siguientes ecuaciones con algún valor propio λ:

$$\begin{aligned} x + y - z &= \lambda x, \\ y \qquad &= \lambda y, \\ x \qquad + z &= \lambda z. \end{aligned}$$

Este sistema es equivalente a

$$\begin{aligned} (1-\lambda)x + y - z &= 0, \\ (1-\lambda)y &= 0, \\ x + (1-\lambda)z &= 0. \end{aligned}$$

Caso 1. $\lambda = 1$. Entonces la segunda ecuación se cumplirá para cualquier valor de y. Póngase $y = 1$; a partir de la primera ecuación obtenemos $z = 1$, y de la tercera ecuación obtenemos $x = 0$. En consecuencia, obtenemos un primer vector propio

$$X^1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Caso 2. $\lambda \neq 1$. Entonces, por la segunda ecuación, debemos tener $y = 0$. Ahora resolvamos el sistema que se forma con la primera y la tercera ecuaciones:

$$(1-\lambda)x - z = 0,$$
$$x + (1-\lambda)z = 0.$$

Si estas ecuaciones fueran independientes, entonces las únicas soluciones serían $x = z = 0$. Éste no puede ser el caso, ya que debe haber un vector propio no nulo que tiene el valor propio dado. En realidad, el lector puede verificar en forma directa que la segunda ecuación es igual a $(\lambda - 1)$ por la primera. En cualquier caso, damos a una de las variables un valor arbitrario y despejamos la otra. Por ejemplo, sea $z = 1$; entonces $x = 1/(1-\lambda)$. Así obtenemos el vector propio

$$X(\lambda) = \begin{pmatrix} 1/(1-\lambda) \\ 0 \\ 1 \end{pmatrix}.$$

Podemos sustituir $\lambda = \lambda_1$ y $\lambda = \lambda_2$ para obtener los vectores propios que tienen los valores propios λ_1 y λ_2, respectivamente.

De esta manera hemos hallado tres vectores propios que tienen valores propios distintos entre sí, a saber,

$$X^1, \qquad X(\lambda_1), \qquad X(\lambda_2).$$

Ejemplo 8. Halle los valores propios y una base para los espacios propios de la matriz

$$\begin{pmatrix} 1 & -1 & 2 \\ -2 & 1 & 3 \\ 1 & -1 & 1 \end{pmatrix}.$$

El polinomio característico es

$$\begin{vmatrix} t-1 & 1 & -2 \\ 2 & t-1 & -3 \\ -1 & 1 & t-1 \end{vmatrix} = (t-1)^3 - (t-1) - 1.$$

Los valores propios son las raíces de esta ecuación cúbica. En general no es fácil hallar tales raíces, y éste es el caso en el ejemplo presente. Sea $u = t - 1$. En términos de u, el polinomio se puede escribir de la manera siguiente:

$$Q(u) = u^3 - u - 1.$$

Por la aritmética tenemos que las únicas raíces racionales deben ser enteras y deben dividir a 1, de manera que las únicas raíces racionales posibles son ± 1, que no son raíces. En consecuencia, no hay valor propio racional. Sin embargo, una ecuación cúbica tiene la forma general que se muestra en la figura 1.

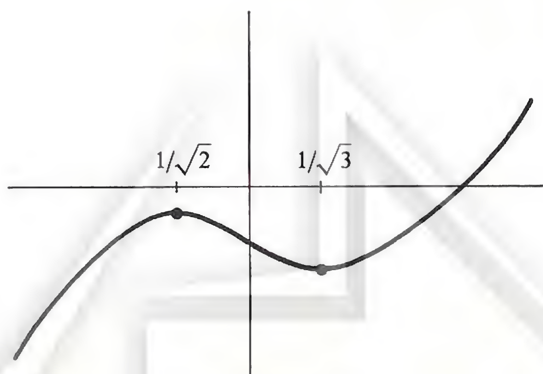


Figura 1

Esto significa que hay al menos una raíz real. Si el lector sabe cálculo, entonces tiene las herramientas que le permitirán determinar el máximo y el mínimo relativos, y hallará que la función $u^3 - u - 1$ tiene su mínimo relativo en $u = 1/\sqrt{3}$, y que $Q(1/\sqrt{3})$ es positivo. Por tanto, sólo hay una raíz real. Las otras dos raíces son complejas. Esto es lo más que podemos hacer con los medios con que contamos. En cualquier caso, damos a estas raíces un nombre y llamamos a los valores propios

$$\lambda_1, \lambda_2, \lambda_3.$$

Todos son distintos entre sí.

Sin embargo, podemos hallar vectores propios en términos de los valores propios. Sea

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

un vector no nulo. Entonces X es un vector propio si, y sólo si, $AX = \lambda X$, esto es:

$$\begin{aligned} x - y + 2z &= \lambda x, \\ -2x + y + 3z &= \lambda y, \\ x - y + z &= \lambda z. \end{aligned}$$

Este sistema de ecuaciones es equivalente a

$$\begin{aligned} (1-\lambda)x - y + 2z &= 0, \\ -2x + (1-\lambda)y + 3z &= 0, \\ x - y + (1-\lambda)z &= 0. \end{aligned}$$

Demos a z un valor arbitrario, digamos $z = 1$, y resolvamos para x y y usando las dos primeras ecuaciones. Así, debemos resolver:

$$\begin{aligned} (\lambda - 1)x + y &= 2, \\ 2x + (\lambda - 1)y &= 3. \end{aligned}$$

Multiplique la primera ecuación por 2, la segunda por $(\lambda - 1)$ y reste. Luego podemos despejar y para obtener

$$y(\lambda) = \frac{3(\lambda - 1) - 4}{(\lambda - 1)^2 - 2}.$$

Por la primera ecuación hallamos que

$$x(\lambda) = \frac{2 - y}{\lambda - 1}.$$

Por tanto, los vectores propios son

$$X(\lambda_1) = \begin{pmatrix} x(\lambda_1) \\ y(\lambda_1) \\ 1 \end{pmatrix}, \qquad X(\lambda_2) = \begin{pmatrix} x(\lambda_2) \\ y(\lambda_2) \\ 1 \end{pmatrix}, \qquad X(\lambda_3) = \begin{pmatrix} x(\lambda_3) \\ y(\lambda_3) \\ 1 \end{pmatrix},$$

donde λ_1, λ_2 y λ_3 son los tres valores propios. Ésta es una respuesta explícita en la medida en que el lector sea capaz de determinar estos valores propios. El lector puede recurrir a una calculadora o una computadora para obtener aproximaciones a λ_1, λ_2 y λ_3 que le darán las correspondientes aproximaciones a los tres vectores propios. Observe que en esta parte hemos hallado los vectores propios complejos. Sea λ_1 el valor propio real (hemos visto que aquí sólo hay uno). Entonces, a partir de las fórmulas para las coordenadas de $X(\lambda)$, vemos que $y(\lambda)$ o $x(\lambda)$ serán reales si, y sólo si, λ es real. Por tanto, sólo hay un vector propio real, a saber, $X(\lambda_1)$. Los otros dos vectores propios son complejos. Cada vector propio es una base para el correspondiente espacio propio.

Teorema 2.3. *Sean A y B dos matrices de $n \times n$ y suponga que B es invertible. Entonces el polinomio característico de A es igual al polinomio característico de $B^{-1}AB$.*

Demostración. Por definición y por las propiedades del determinante,

$$\begin{aligned}\operatorname{Det}(tI - A) &= \operatorname{Det}(B^{-1}(tI - A)B) = \operatorname{Det}(tB^{-1}B - B^{-1}AB) \\ &= \operatorname{Det}(tI - B^{-1}AB).\end{aligned}$$

Esto prueba lo que queríamos.

Ejercicios VIII, §2

1. Sea A una matriz diagonal,

$$A = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}.$$

(a) ¿Cuál es el polinomio característico de A?
(b) ¿Cuáles son sus valores propios?

2. Sea A una matriz triangular,

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

¿Cuál es el polinomio característico de A y cuáles son sus valores propios?

Encuentre el polinomio característico, los valores propios y las bases para los espacios propios de las siguientes matrices.

3. (a) $\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}$

(c) $\begin{pmatrix} -2 & -7 \\ 1 & 2 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$

4. (a) $\begin{pmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}$

(c) $\begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 3 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ -1 & 1 & 4 \end{pmatrix}$

5. Encuentre los valores propios y los vectores propios de las siguientes matrices. Demuestre que los vectores propios forman un espacio de dimensión 1.

(a) $\begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$ (d) $\begin{pmatrix} 2 & -3 \\ 1 & -1 \end{pmatrix}$

6. Encuentre los valores propios y los vectores propios de las siguientes matrices. Demuestre que los vectores propios forman un espacio de dimensión 1.

(a) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

7. Halle los valores propios y una base para los espacios propios de las siguientes matrices.

(a) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} -1 & 0 & 1 \\ -1 & 3 & 0 \\ -4 & 13 & -1 \end{pmatrix}$

8. Halle los valores propios y una base para los espacios propios de las siguientes matrices.

(a) $\begin{pmatrix} 2 & 4 \\ 5 & 3 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$ (c) $\begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix}$

(d) $\begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{pmatrix}$ (e) $\begin{pmatrix} 3 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & -1 \end{pmatrix}$ (f) $\begin{pmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{pmatrix}$

9. Sea V un espacio vectorial de dimensión n y suponga que el polinomio característico de una aplicación lineal $A: V \to V$ tiene n raíces distintas. Demuestre que V tiene una base que consta de vectores propios de A.

10. Sea A una matriz cuadrada. Demuestre que los valores propios de ${}^tA$ son los mismos que los de A.

11. Sea A una matriz invertible. Si λ es un valor propio de A, demuestre que $\lambda \neq 0$ y que λ^{-1} es un valor propio de A^{-1}.

12. Sea V el espacio generado sobre $\mathbf{R}$ por las funciones $\operatorname{sen} t$ y $\cos t$. ¿Tiene la derivada (considerada como una aplicación lineal de V en sí mismo) algunos vectores propios no nulos en V? Si así es, ¿cuáles son?

13. Denote con D la derivada que consideramos como aplicación lineal sobre el espacio de las funciones diferenciables. Sea k un entero $\neq 0$. Demuestre que las funciones $\operatorname{sen} kx$ y $\cos kx$ son valores propios en D^2. ¿Cuáles son los valores propios?

14. Sea $A: V \to V$ una aplicación lineal de V en sí mismo, y sea $\{v_1, \ldots, v_n\}$ una base de V que consta de vectores propios que tienen distintos valores propios $c_1, \ldots, c_n$. Demuestre que cualquier vector propio v de A en V es un múltiplo escalar de algún v_i.

15. Sean A y B matrices cuadradas del mismo tamaño. Muestre que los valores propios de AB son los mismos que los valores propios de BA.

VIII, §3. Valores propios y vectores propios de matrices simétricas

Daremos dos pruebas del siguiente teorema.

Teorema 3.1. *Sea A una matriz real simétrica de $n \times n$. Entonces existe un vector propio no nulo para A.*

Una de las pruebas usará números complejos y la otra prueba usará cálculo. Comencemos con la prueba que usa cálculo.

Defina la función

$$f(X) = {}^tXAX \qquad \text{para} \qquad X \in \mathbf{R}^n.$$

Tal función f se conoce como **forma cuadrática asociada con A**. Si ${}^tX = (x_1, \ldots, x_n)$ se escribe en términos de coordenadas y $A = (a_{ij})$, entonces

$$f(X) = \textstyle\sum_{i,j=1}^{n} a_{ij} x_i x_j.$$

Ejemplo. Sea

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}$$

Sea ${}^tX = (x, y)$. Entonces

$${}^tXAX = (x, y) \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 3x^2 - 2xy + 2y^2.$$

En forma más general, sea

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}.$$

Entonces

$$(x, y)\begin{pmatrix} a & b \\ b & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2bxy + dy^2.$$

Ejemplo. Suponga que nos dan una expresión cuadrática

$$f(x, y) = 3x^2 + 5xy - 4y^2.$$

Entonces ésta es la forma cuadrática asociada con la matriz simétrica

$$A = \begin{pmatrix} 3 & \frac{5}{2} \\ \frac{5}{2} & -4 \end{pmatrix}.$$

En muchas aplicaciones se quiere hallar un máximo para dicha función f sobre la esfera unitaria. Recuerde que la **esfera unitaria** es el conjunto de todos los elementos X tales que $\|X\| = 1$, donde $\|X\| = \sqrt{X \cdot X}$. En los cursos de análisis se demuestra que una función continua f como la anterior necesariamente tiene un máximo sobre la esfera. Un **máximo** sobre la esfera unitaria es un punto P tal que $\|P\| = 1$, y

$$f(P) \geq f(X) \qquad \text{para todo } X \text{ con } \|X\| = 1.$$

El siguiente teorema relaciona este problema con el de hallar vectores propios.

Teorema 3.2. *Sea A una matriz simétrica real y sea $f(X) = {}^tXAX$ la forma cuadrática asociada. Sea P un punto sobre la esfera unitaria tal que $f(P)$ es un máximo para f sobre la esfera. Entonces P es un vector propio para A. En otras palabras, existe un número λ tal que $AP = \lambda P$.*

Demostración. Sea W el subespacio de $\mathbf{R}^n$ ortogonal a P, esto es, $W = P^{\perp}$. Entonces $\dim W = n - 1$. Para cualquier elemento $w \in W$, $\|w\| = 1$, defina la curva

$$C(t) = (\cos t)P + (\operatorname{sen} t)w.$$

Las direcciones de los vectores unitarios $w \in W$ son las direcciones de las tangentes a la esfera en el punto P, como se muestra en la figura.

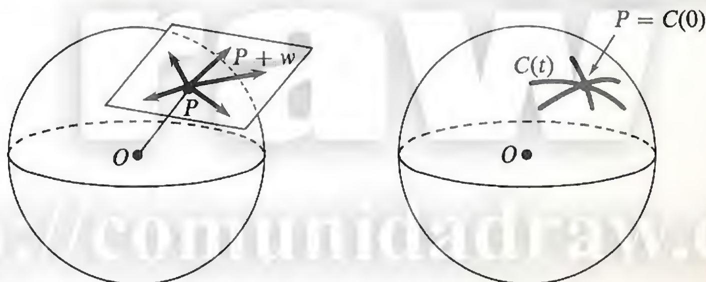


Figura 2

La curva $C(t)$ está sobre la esfera debido a que $\|C(t)\| = 1$, como se puede verificar fácilmente al considerar el producto interior $C(t) \cdot C(t)$ y al usar la

hipótesis de que $P \cdot w = 0$. Además, $C(0) = P$, de manera que $C(t)$ es una curva sobre la esfera que pasa por P. También tenemos la derivada siguiente

$$C'(t) = (-\operatorname{sen} t)P + (\cos t)w,$$

y, por tanto, $C'(0) = w$. Por ello la dirección de la curva es la dirección de w, y es tangente a la esfera en el punto P porque $w \cdot P = 0$. Consideremos la función

$$g(t) = f\big(C(t)\big) = C(t) \cdot AC(t).$$

Al usar coordenadas y la regla para la derivada de un producto, la cual se aplica en este caso (como el lector sabe por sus estudios de cálculo), se hallará la siguiente derivada:

$$\begin{aligned} g'(t) &= C'(t) \cdot AC(t) + C(t) \cdot AC'(t) \\ &= 2C'(t) \cdot AC(t), \end{aligned}$$

ya que A es simétrica. Como $f(P)$ es un máximo y $g(0) = f(P)$, se infiere que $g'(0) = 0$. Entonces obtenemos

$$O = g'(0) = 2C'(0) \cdot AC(0) = 2w \cdot AP.$$

En consecuencia, AP es perpendicular a W para todo $w \in W$. Pero $W^{\perp}$ es el espacio de dimensión 1 generado por P. Por tanto, existe un número λ tal que $AP = \lambda P$, lo que prueba el teorema.

Corolario 3.3. *El valor máximo de f sobre la esfera unitaria es igual al valor propio más grande de A.*

Demostración. Sea λ cualquier valor propio y sea P un vector propio sobre la esfera unitaria, de manera que $\|P\| = 1$. Entonces

$$f(P) = {}^tPAP = {}^tP\lambda P = \lambda {}^tPP = \lambda.$$

Por tanto, el valor de f en un vector propio sobre la esfera unitaria es igual al valor propio. El Teorema 3.2 nos dice que el máximo de f sobre la esfera unitaria se alcanza en un vector propio, por lo que el máximo de f sobre la esfera unitaria es igual al mayor valor propio, tal como se afirmó.

Ejemplo. Sea $f(x, y) = 2x^2 - 3xy + y^2$. Sea A la matriz simétrica asociada con f. Encuentre los vectores propios de A sobre el círculo unitario y halle el máximo de f sobre el círculo unitario.

Primero observamos que f es la forma cuadrática asociada con la matriz

$$A = \begin{pmatrix} 2 & -\frac{3}{2} \\ -\frac{3}{2} & 1 \end{pmatrix}.$$

Por el Teorema 3.2, se debe alcanzar un máximo en un vector propio, de manera que primero determinamos los valores propios y los vectores propios.

El polinomio característico es el determinante

$$\begin{vmatrix} t-2 & \frac{3}{2} \\ \frac{3}{2} & t-1 \end{vmatrix} = t^2 - 3t - \tfrac{1}{4}.$$

Luego, los valores propios son

$$\lambda = \frac{3 \pm \sqrt{10}}{2}.$$

Respecto a los vectores propios, debemos resolver el siguiente sistema:

$$\begin{aligned} 2x - \tfrac{3}{2}y &= \lambda x, \\ -\tfrac{3}{2}x - \quad y &= \lambda y. \end{aligned}$$

Al hacer $x = 1$ se obtienen los posibles vectores propios

$$X(\lambda) = \begin{pmatrix} 1 \\ \frac{2}{3}(2-\lambda) \end{pmatrix}.$$

Por tanto, hay dos de tales vectores propios, salvo múltiplos escalares no nulos. Los vectores propios que están sobre el círculo unitario son, por consiguiente,

$$P(\lambda) = \frac{X(\lambda)}{\|X(\lambda)\|} \qquad \text{donde} \qquad \lambda = \frac{3+\sqrt{10}}{2} \qquad \text{y} \qquad \lambda = \frac{3-\sqrt{10}}{2}.$$

Por el Corolario 3.3, el máximo es el punto que tiene el valor propio más grande y, en consecuencia, debe ser el punto

$$P(\lambda) \qquad \text{donde} \qquad \lambda = \frac{3+\sqrt{10}}{2}.$$

El valor máximo de f sobre el círculo unitario es $(3 + \sqrt{10})/2$.

Asimismo, el valor mínimo de f sobre el círculo unitario es $(3 - \sqrt{10})/2$.

Ahora usaremos los números complejos $\mathbf{C}$ para la segunda prueba. Una propiedad fundamental de los números complejos es que todo polinomio no constante con coeficientes complejos tiene una raíz (un cero) en los números complejos. Por consiguiente, el polinomio característico de A tiene una raíz compleja λ que, a priori, es un valor propio complejo, que tiene un vector propio complejo.

Teorema 3.4. *Sea A una matriz simétrica real y sea λ un valor propio en $\mathbf{C}$. Entonces λ es real. Si $Z \neq O$ es un vector propio complejo que tiene un valor propio λ y $Z = X + iY$, donde $X, Y \in \mathbf{R}^n$, entonces tanto X como Y son vectores propios reales de A que tienen valor propio λ, y X o $Y \neq O$.*

Demostración. Sea $Z = {}^t(z_1, \ldots, z_n)$ con coordenadas complejas z_i. Entonces

$$Z \cdot \overline{Z} = \overline{Z} \cdot Z = {}^t\overline{Z}Z = \overline{z}_1 z_1 + \cdots + \overline{z}_n z_n = |z_1|^2 + \cdots + |z_n|^2 > 0.$$

Por hipótesis, tenemos que $AZ = \lambda Z$. Entonces

$${}^t\overline{Z}AZ = {}^t\overline{Z}\lambda Z = \lambda {}^t\overline{Z}Z.$$

La transpuesta de una matriz de 1×1 es igual a sí misma, de manera que también obtenemos

$${}^tZ{}^tA\overline{Z} = {}^tZA\overline{Z} = \lambda {}^t\overline{Z}Z.$$

Pero $\overline{AZ} = \overline{A}\,\overline{Z} = A\overline{Z}$, y $\overline{AZ} = \overline{\lambda Z} = \overline{\lambda}\,\overline{Z}$. Por consiguiente,

$$\lambda {}^t\overline{Z}Z = \overline{\lambda}{}^tZ\overline{Z}.$$

Como ${}^tZ\overline{Z} \neq 0$, se infiere que $\lambda = \overline{\lambda}$, de manera que λ es real.

Ahora, a partir de $AZ = \lambda Z$, obtenemos

$$AX + iAY = \lambda X + i\lambda Y,$$

y, como A, X y Y son reales, se infiere que $AX = \lambda X$ y $AY = \lambda Y$. Esto prueba el teorema.

Ejercicios VIII, §3

1. Halle los valores propios de las siguientes matrices y el valor máximo de las formas cuadráticas asociadas sobre el círculo unitario.

(a) $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$

2. La misma pregunta que en el Ejercicio 1, excepto que hay que hallar el máximo en la esfera unitaria.

(a) $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$

3. Halle el máximo y el mínimo de la función

$$f(x, y) = 3x^2 + 5xy - 4y^2$$

sobre el círculo unitario.

VIII, §4. Diagonalización de una aplicación lineal simétrica

En esta sección aplicaremos la existencia de vectores propios conforme a lo probado en la sección §3. Como lo haremos por inducción, en lugar de trabajar con $\mathbf{R}^n$ tenemos que comenzar con una formulación que trate con un espacio vectorial en el que todavía no se hayan escogido coordenadas.

Por tanto, sea V un espacio vectorial de dimensión n sobre $\mathbf{R}$, que tiene un producto escalar difinitivamente positivo $\langle v, w\rangle$ para v, $w \in V$. Sea

$$A: V \to V$$

una aplicación lineal. Diremos que A es **simétrica (respecto al producto escalar)** si tenemos la relación

$$\langle Av, w\rangle = \langle v, Aw\rangle$$

para todo v y $w \in V$.

Ejemplo. Suponga que $V = \mathbf{R}^n$ y que el producto escalar es el producto interior usual entre vectores. Una aplicación lineal A está representada por una matriz y usamos la misma letra A para esta matriz. Entonces, para todo v, $w \in \mathbf{R}^n$, tenemos

$$\langle Av, w\rangle = {}^t wAv$$

si consideramos a v y w como vectores columna. Como ${}^t wAv$ es una matriz de 1×1, es igual a su transpuesta, de modo que tenemos la fórmula

$${}^t wAv = {}^t v {}^t Aw,$$

o, en términos de la notación $\langle\ ,\ \rangle$,

$$\langle Av, w\rangle = \langle v, {}^t Aw\rangle.$$

La condición de que A sea simétrica como aplicación lineal con respecto al producto escalar es, por definición, $\langle Av, w\rangle = \langle v, Aw\rangle$ o, en términos de multiplicación de matrices,

$${}^t wAv = {}^t vAw$$

Al comparar con la fórmula anterior, esto significa que $A = {}^t A$, por lo que hallamos:

Sea A una matriz de $n \times n$ y L_A la aplicación lineal asociada sobre $\mathbf{R}^n$. Entonces L_A es simétrica con respecto al producto escalar si, y sólo si, A es una matriz simétrica.

Si V es un espacio vectorial general de dimensión n con un producto escalar definitivamente positivo, y $A: V \to V$ es una aplicación lineal de V en sí mismo, entonces existe una aplicación lineal única

$${}^t A: V \to V$$

que satisface la fórmula

$$\boxed{\langle Av, w\rangle = \langle v, {}^t Aw\rangle.}$$

Acabamos de ver esto cuando $V = \mathbf{R}^n$. En general, escogemos una base ortonormal de V, que nos permite identificar a V con $\mathbf{R}^n$ e identificar al producto escalar con el producto interior ordinario. Entonces ${}^t A$ (considerada como aplicación lineal) coincide con la transpuesta de la matriz que representa a A.

Ahora podemos decir que una aplicación lineal $A: V \to V$ es simétrica si, y sólo si, es igual a su propia transpuesta. Cuando V se identifica con $\mathbf{R}^n$ utilizando una base ortonormal, significa que la matriz que representa a A es igual a su transpuesta; en otras palabras, la matriz es simétrica.

Podemos reformular el Teorema 3.1 de la manera siguiente:

Sea V un espacio vectorial de dimensión finita con un producto escalar definitivamente positivo. Sea $A: V \to V$ una aplicación lineal simétrica. Entonces A tiene un vector propio no nulo.

Sea W un subespacio de V y sea $A: V \to V$ una aplicación lineal simétrica. Decimos que W es **estable bajo** A si $A(W) \subset W$, esto es, para todo $u \in W$ tenemos que $Au \in W$.

Observemos que, si W es estable bajo A, entonces su complemento ortogonal $W^{\perp}$ también es estable bajo A.

Demostración. Sea $w \in W^{\perp}$. Entonces, para todo $u \in W$, tenemos que

$$\langle Aw, u \rangle = \langle w, Au \rangle = 0$$

debido a que $Au \in W$ y $w \in W^{\perp}$. En consecuencia, $Aw \in W^{\perp}$, lo que prueba el aserto.

Teorema 4.1. *Sea V un espacio vectorial de dimensión finita sobre los números reales, de dimensión $n > 0$ y con un producto escalar definitivamente positivo. Sea*

$$A: V \to V$$

una aplicación lineal, simétrica con respecto al producto escalar. Entonces V tiene una base ortonormal que consta de vectores propios.

Demostración. Por el Teorema 3.1, existe un vector propio no nulo P para A. Sea W el espacio de una dimensión generado por P. Entonces W es estable bajo A. Por la observación anterior, $W^{\perp}$ también es estable bajo A y es un espacio vectorial de dimensión $n-1$. Podemos considerar entonces que A da una aplicación lineal simétrica de $W^{\perp}$ en sí mismo. Luego podemos repetir el procedimiento. Ponemos $P = P_1$ y, por indución, podemos encontrar una base $\{P_2 \ldots, P_n\}$ de $W^{\perp}$ que consta de vectores propios. Entonces

$$\{P_1, P_2, \ldots, P_n\}$$

es una base ortogonal de V que consta de vectores propios. Dividamos cada vector entre su norma para obtener una base ortonormal, como se deseaba.

Si $\{e_1, \ldots, e_n\}$ es una base ortonormal de V tal que cada e_i es un vector propio, entonces la matriz de A con respecto a esta base es diagonal y los elementos diagonales son precisamente los valores propios:

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

En esta sencilla representación, el efecto de A resulta mucho más claro que cuando A se representa mediante una matriz más complicada con respecto a otra base.

Ejemplo. Veamos una aplicación a las ecuaciones diferenciales lineales. Sea A una matriz real simétrica de $n \times n$. Queremos encontrar soluciones en $\mathbf{R}^n$ de la siguiente ecuación diferencial

$$\frac{dX(t)}{dt} = AX(t),$$

donde

$$X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$

está dada en términos de coordenadas que son funciones de t, y

$$\frac{dX(t)}{dt} = \begin{pmatrix} dx_1/dt \\ \vdots \\ dx_n/dt \end{pmatrix}.$$

La escritura de esta ecuación en términos de coordenadas arbitrarias es confusa. Así que, por el momento, olvidemos las coordenadas, y consideremos $\mathbf{R}^n$ como un espacio vectorial de dimensión n con un producto escalar definitivamente positivo. Escojamos una base ortonormal de V (usualmente diferente de la base original) que consta de vectores propios de A. Ahora, *con respecto a esta nueva base*, podemos identificar a V con $\mathbf{R}^n$ con las nuevas coordenadas, las que denotamos con $y_1, \ldots, y_n$. Con respecto a estas nuevas coordenadas, la matriz de la aplicación lineal L_A es la siguiente

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

donde $\lambda_1, \ldots, \lambda_n$ son los valores propios. Pero en términos de estas coordenadas, más convenientes, nuestra ecuación diferencial simplemente se ve de la siguiente manera:

$$\frac{dy_1}{dt} = \lambda_1 y_1, \quad \ldots, \quad \frac{dy_n}{dt} = \lambda_n y_n.$$

Por tanto, la solución más general es de la forma

$$y_i(t) = c_i e^{\lambda_i t} \qquad \text{con alguna constante } c_i.$$

La moraleja de este ejemplo es que no se debería seleccionar una base con precipitación, y se debería usar, tan a menudo como sea posible, una notación sin coordenadas, hasta que se haga imperativo realizar una selección de coordenadas para simplificar la solución de un problema.

Ejercicios VIII, §4

1. Suponga que A es una matriz diagonal de $n \times n$. Para cualquier $X \in \mathbf{R}^n$, ¿a qué es igual ${}^tXAX$ en términos de las coordenadas de X y los elementos diagonales de A?

2. Sea

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

una matriz diagonal con $\lambda_1 \geq 0 \ldots, \lambda_n \geq 0$. Demuestre que existe una matriz diagonal B de $n \times n$ tal que $B^2 = A$.

3. Sea V un espacio vectorial de dimensión finita con un producto escalar definitivamente positivo. Sea $A: V \to V$ una aplicación lineal simétrica. Decimos que A es **definitivamente positiva** si $\langle Av, v\rangle > 0$ para todo $v \in V$, y $v \neq 0$. Pruebe que
   (a) si A es definitivamente positiva, entonces todos los valores propios son > 0.
   (b) si A es definitivamente positiva, entonces existe una aplicación lineal simétrica B tal que $B^2 = A$ y $BA = AB$. ¿Cuáles son los valores propios de B? [*Sugerencia:* Use una base de V que conste de vectores propios.]

4. Decimos que A es **semidefinitivamente positiva** si $\langle Av, v\rangle \geq 0$ para todo $v \in V$. Pruebe los análogos de (a) y (b) del Ejercicio 3 cuando se supone que A sólo es semidefinitiva. Por tanto, los valores propios son ≥ 0, y existe una aplicación lineal simétrica B tal que $B^2 = A$.

5. Suponga que A es simétrica y definitivamente positiva. Demuestre que A^2 y A^{-1} son simétricas y definitivamente positivas.

6. Sea $U: \mathbf{R}^n \to \mathbf{R}^n$ una aplicación lineal, y denote con $\langle\ ,\ \rangle$ al producto escalar (interior) usual. Demuestre que las siguientes condiciones son equivalentes:
   (i) $\|Uv\| = \|v\|$ para todo $v \in \mathbf{R}^n$.
   (ii) $\langle Uv, Uw\rangle = \langle v, w\rangle$ para todo $v, w \in \mathbf{R}^n$.
   (iii) U es invertible y ${}^tU = U^{-1}$.
   [*Sugerencia:* Para (ii), use la identidad siguiente:

$$\langle (v+w), (v+w)\rangle - \langle (v-w), (v-w)\rangle = 4\langle v, w\rangle,$$

   y en forma análoga, con una U frente a cada vector.] Cuando U satisface cualquiera de estas condiciones (y, en consecuencia, todas ellas), se dice que U es **unitaria**. La primera condición dice que U preserva la norma y la segunda dice que U preserva el producto escalar.

7. Sea $A: \mathbf{R}^n \to \mathbf{R}^n$ una aplicación lineal invertible.
   (i) Demuestre que ${}^tAA$ es simétrica y definitivamente positiva.
   (ii) Por el Ejercicio 3b, existe una matriz B simétrica y definitivamente positiva tal que $B^2 = {}^tAA$. Sea $U = AB^{-1}$; demuestre que U es unitaria.
   (iii) Demuestre que $A = UB$.

8. Sea B simétrica y definitivamente positiva, y unitaria además. Demuestre que $B = I$.

Apéndice: Números complejos

Los **números complejos** $\mathbf{C}$ son un conjunto de objetos que se pueden sumar y multiplicar, de tal manera que la suma y el producto de dos números complejos también son números complejos y satisfacen las siguientes condiciones

(1) Todo número real es un número complejo y, si α y β son números reales, entonces su suma y su producto como números complejos son iguales a su suma y su producto como números reales.

(2) Existe un número complejo, denotado con i tal que $i^2 = -1$.

(3) Todo número complejo se puede escribir en forma única en la forma $a+bi$, donde a y b son números reales.

(4) Se cumplen las leyes ordinarias de la aritmética concernientes a la adición y la multiplicación. Damos la lista de estas leyes:

Si α, β y γ son números complejos, entonces

$$(\alpha+\beta)+\gamma = \alpha+(\beta+\gamma) \qquad \text{y} \qquad (\alpha\beta)\gamma = \alpha(\beta\gamma).$$

Tenemos que $\alpha(\beta+\gamma) = \alpha\beta+\alpha\gamma$, y $(\beta+\gamma)\alpha = \beta\alpha+\gamma\alpha$.

Tenemos que $\alpha\beta = \beta\alpha$ y $\alpha+\beta = \beta+\alpha$.

Si 1 es el número real uno, entonces $1\alpha = \alpha$.

Si 0 es el número real cero, entonces $0\alpha = 0$.

Tenemos que $\alpha+(-1)\alpha = 0$.

Ahora derivaremos consecuencias de estas propiedades. Si escribimos

$$\alpha = a_1 + a_2 i \qquad \text{y} \qquad \beta = b_1 + b_1 i,$$

entonces

$$\alpha+\beta = a_1 + a_2 i + b_1 + b_2 i = a_1 + b_1 + (a_2 + b_2)i.$$

Si llamamos a a_1 la **parte real**, o componente real de α y a a_2 su **parte imaginaria**, o componente imaginaria, entonces vemos que la adición se lleva a cabo componente a componente. Las partes real e imaginaria de α se denotan con $\mathrm{Re}(\alpha)$ e $\mathrm{Im}(\alpha)$, respectivamente.

Tenemos que

$$\alpha\beta = (a_1 + a_2 i)(b_1 + b_2 i) = a_1 b_1 - a_2 b_2 + (a_2 b_1 + a_1 b_2)i.$$

Sea $\alpha = a + bi$ un número complejo, con a y b reales. Definimos

$$\overline{\alpha} = a - bi$$

y decimos que $\overline{\alpha}$ es el conjugado complejo, o simplemente **conjugado**, de α. Entonces

$$\alpha\overline{\alpha} = a^2 + b^2.$$

Si $\alpha = a + bi$ es $\neq 0$, y si hacemos

$$\lambda = \frac{\overline{\alpha}}{a^2 + b^2},$$

entonces $\alpha\lambda = \lambda\alpha = 1$, como de inmediato se puede ver. El número λ anterior se conoce como **inverso** de α y se denota con α^{-1} ó $1/\alpha$. Observemos que es el único número complejo z tal que $z\alpha = 1$, ya que, si se satisface esta ecuación, la multiplicamos por λ a la derecha para hallar $z = \lambda$. Si α y β son números complejos, a menudo escribimos β/α en lugar de $\alpha^{-1}\beta$ o de $\beta\alpha^{-1}$. Vemos que podemos dividir entre números complejos $\neq 0$.

Tenemos las siguientes reglas:

$$\overline{\alpha\beta} = \overline{\alpha}\overline{\beta}, \qquad \overline{\alpha+\beta} = \overline{\alpha}+\overline{\beta}, \qquad \overline{\overline{\alpha}} = \alpha.$$

Éstas se infieren de inmediato a partir de la definición de adición y multiplicación.

Definimos el **valor absoluto** de un número complejo $\alpha = a + bi$ como

$$|\alpha| = \sqrt{a^2 + b^2}.$$

Si imaginamos α como un punto en el plano (a, b), entonces $|\alpha|$ es la longitud del segmento de recta que va del origen a α. En términos del valor absoluto, podemos escribir

$$\alpha^{-1} = \frac{\overline{\alpha}}{|\alpha|^2}$$

siempre que $\alpha \neq 0$. En efecto, observemos que $|\alpha|^2 = \alpha\overline{\alpha}$. Advierta también que $|\alpha| = |\overline{\alpha}|$.

El valor absoluto satisface propiedades análogas a las que satisface el valor absoluto de números reales:

$$|\alpha| \geq 0 \text{ y } = 0 \text{ si, y sólo si, } \alpha = 0.$$

$$|\alpha\beta| = |\alpha|\,|\beta|$$

$$|\alpha + \beta| \leq |\alpha| + |\beta|.$$

El primer aserto es obvio. Por lo que se refiere al segundo, tenemos

$$|\alpha\beta|^2 = \alpha\beta\overline{\alpha}\overline{\beta} = |\alpha|^2|\beta|^2.$$

Al extraer la raíz cuadrada, concluimos que $|\alpha|\,|\beta| = |\alpha\beta|$. Después, tenemos

$$|\alpha + \beta|^2 = (\alpha + \beta)\overline{(\alpha + \beta)} = (\alpha + \beta)(\overline{\alpha} + \overline{\beta})$$

ya que $\alpha\overline{\beta} = \overline{\beta\overline{\alpha}}$. Sin embargo, tenemos que

$$2\mathrm{Re}(\beta\overline{\alpha}) \leq 2|\beta\overline{\alpha}|$$

debido a que la parte real de un número complejo es $\leq$ que su valor absoluto. Por tanto,

$$\begin{aligned} |\alpha + \beta|^2 &\leq |\alpha|^2 + 2|\beta\overline{\alpha}| + |\beta|^2 \\ &\leq |\alpha|^2 + 2|\beta|\,|\alpha| + |\beta|^2 \\ &= (|\alpha| + |\beta|)^2. \end{aligned}$$

Al extraer la raíz cuadrada se obtiene la propiedad final.

Sea $z = x + iy$ un número complejo $\neq 0$. Entonces $z/|z|$ tiene valor absoluto igual a 1.

La principal ventaja de trabajar con números complejos en lugar de hacerlo con números reales es que todo polinomio no constante con coeficientes complejos tiene una raíz en **C**. Esto se prueba en cursos de análisis más avanzados. Por ejemplo, una ecuación cuadrática

$$ax^2 + bx + c = 0$$

con $a \neq 0$, tiene las raíces

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Si $b^2 - 4ac$ es positivo, entonces las raíces son reales. Si $b^2 - 4ac$ es negativo, entonces las raíces son complejas. La prueba para la fórmula cuadrática emplea sólo la aritmética básica de la adición, la multiplicación y la división. A saber, completamos el cuadrado para ver que

$$ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a} + c = a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a}.$$

Luego resolvemos

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2},$$

extraemos la raíz cuadrada y, por último, obtenemos el valor deseado de x.

Aplicación a los espacios vectoriales

Para definir la noción de espacio vectorial, primero necesitamos la noción de escalares; los únicos hechos que necesitamos acerca de los escalares son los que se refieren a la adición, multiplicación y división entre elementos no nulos. Los números complejos satisfacen todas las operaciones básicas de la aritmética. Por consiguiente, podemos hacer la teoría básica de los espacios vectoriales sobre los números complejos. Tenemos los mismos teoremas sobre combinaciones lineales, matrices, rango por renglones, rango por columnas, dimensión, determinantes, polinomios característicos, valores propios.

La única diferencia básica (y es ligera) aparece cuando trabajamos con el producto interior. Si $Z = (z_1, \ldots, z_n)$ y $W = (w_1, \ldots, w_n)$ son n-tuplas de $\mathbf{C}^n$, entonces su producto interior es, como antes,

$$Z \cdot W = z_1 w_1 + \cdots + z_n w_n.$$

Sin embargo, observe que, aun si $Z \neq O$, $Z \cdot Z$ puede ser igual a 0. Por ejemplo, sea $Z = (1, i)$ en $\mathbf{C}^2$. Entonces

$$Z \cdot Z = 1 + i^2 = 1 - 1 = 0.$$

Por tanto, el producto interior no es definitivamente positivo.

Para remediar esto, se define un producto que se llama **hermitiano** y es casi igual al producto interior, pero contiene un conjunto complejo. Esto es, definimos $\overline{W} = (\overline{w}_1, \ldots, \overline{w}_n)$ y

$$\langle Z, W \rangle = Z \cdot \overline{W} = z_1 \overline{w}_1 + \cdots + z_n \overline{w}_n,$$

de manera que ponemos un *conjugado complejo en las coordenadas de* W. Entonces

$$\langle Z, Z \rangle = Z \cdot \overline{Z} = z_1 \overline{z}_1 + \cdots + z_n \overline{z}_n = |z_1|^2 + \cdots + |z_n|^2.$$

En consecuencia, una vez más, si $Z \neq O$, entonces alguna coordenada $z_i \neq 0$, de manera que la suma que aparece a la derecha es $\neq 0$ y $\langle Z, Z \rangle > 0$.

Si α es un número complejo, por la definición vemos que

$$\langle \alpha Z, W \rangle = \alpha \langle Z, W \rangle \qquad \text{pero} \qquad \langle Z, \alpha W \rangle = \overline{\alpha} \langle Z, W \rangle.$$

Por tanto, en la segunda fórmula aparece un conjunto complejo. Aún tenemos las fórmulas que expresan la aditividad

$$\langle Z_1 + Z_2, W\rangle = \langle Z_1, W\rangle + \langle Z_2, W\rangle$$

y

$$\langle Z, W_1 + W_2\rangle = \langle Z, W_1\rangle + \langle Z, W_2\rangle.$$

Entonces decimos que el producto hermitiano es **lineal** en la primera variable y **antilineal** en la segunda variable. Observe que, en lugar de la conmutatividad del producto hermitiano, tenemos la fórmula

$$\langle Z, W\rangle = \overline{\langle W, Z\rangle}.$$

Si Z y W son vectores reales, entonces el producto hermitiano $\langle Z, W\rangle$ es igual que el producto interior.

Entonces se puede desarrollar el proceso de ortogonalización de Gram-Schmidt igual que antes, pero usando el producto hermitiano en lugar del producto interior.

En la aplicación de este capítulo no necesitamos el producto hermitiano. Todo lo que necesitamos fue que una matriz compleja A de $n \times n$ tuviera un valor propio y que los valores propios fueran las raíces del polinomio característico

$$\det(tI - A).$$

Tal como se mencionó antes, un polinomio no constante con coeficientes complejos siempre tiene una raíz en los números complejos, de manera que A siempre tiene un valor propio en $\mathbf{C}$. En el texto demostramos que, cuando A es real y simétrica, tales valores propios, de hecho, deben ser reales.

Respuestas a los ejercicios

I, §1, pág. 8

	$A+B$	$A-B$	$3A$	$-2B$
1.	$(1,0)$	$(3,-2)$	$(6,-3)$	$(2,-2)$
2.	$(-1,7)$	$(-1,-1)$	$(-3,9)$	$(0,-8)$
3.	$(1,0,6)$	$(3,-2,4)$	$(6,-3,15)$	$(2,-2,-2)$
4.	$(-2,1,-1)$	$(0,-5,7)$	$(-3,-6,9)$	$(2,-6,8)$
5.	$(3\pi,0,6)$	$(-\pi,6,-8)$	$(3\pi,9,-3)$	$(-4\pi,6,-14)$
6.	$(15+\pi,1,3)$	$(15-\pi,-5,5)$	$(45,-6,12)$	$(-2\pi,-6,2)$

I, §2, pág. 12

1. No **2.** Sí **3.** No **4.** Sí **5.** No **6.** Sí **7.** Sí **8.** No

I, §3, pág. 15

1. (a) 5 (b) 10 (c) 30 (d) 14 (e) π^2+10 (f) 245

2. (a) -3 (b) 12 (c) 2 (d) -17 (e) $2\pi^2-16$ (f) $15\pi-10$

4. (b) y (d)

I, §4, pág. 28

1. (a) $\sqrt{5}$ (b) $\sqrt{10}$ (c) $\sqrt{30}$ (d) $\sqrt{14}$ (e) $\sqrt{10+\pi^2}$ (f) $\sqrt{245}$

2. (a) $\sqrt{2}$ (b) 4 (c) $\sqrt{3}$ (d) $\sqrt{26}$ (e) $\sqrt{58+4\pi^2}$ (f) $\sqrt{10+\pi^2}$

3. (a) $(\frac{3}{2},-\frac{3}{2})$ (b) $(0,3)$ (c) $(-\frac{2}{3},\frac{2}{3},\frac{2}{3})$ (d) $(\frac{17}{26},-\frac{51}{26},\frac{34}{13})$

(e) $\dfrac{\pi^2-8}{2\pi^2+29}(2\pi,-3,7)$ (f) $\dfrac{15\pi-10}{10+\pi^2}(\pi,3,-1)$

4. (a) $(-\frac{6}{5},\frac{3}{5})$ (b) $(-\frac{6}{5},\frac{18}{5})$ (c) $(\frac{2}{15},-\frac{1}{15},\frac{1}{3})$ (d) $-\frac{17}{14}(-1,-2,3)$

(e) $\dfrac{2\pi^2-16}{\pi^2+10}(\pi,3,-1)$ (f) $\dfrac{3\pi-2}{49}(15,-2,4)$

5. (a) $\dfrac{-1}{\sqrt{5}\sqrt{34}}$ (b) $\dfrac{-2}{\sqrt{5}}$ (c) $\dfrac{10}{\sqrt{14}\sqrt{35}}$ (d) $\dfrac{13}{\sqrt{21}\sqrt{11}}$ (e) $\dfrac{-1}{\sqrt{12}}$

6. (a) $\dfrac{35}{\sqrt{41\cdot 35}},\dfrac{6}{\sqrt{41\cdot 6}},0$ (b) $\dfrac{1}{\sqrt{17\cdot 26}},\dfrac{16}{\sqrt{41\cdot 17}},\dfrac{25}{\sqrt{26\cdot 41}}$

7. Multipliquemos escalarmente la suma

$$c_1A_1+\cdots+c_rA_r=O$$

por A_i. Hallamos que

$$c_1A_1\cdot A_i+\cdots+c_iA_i\cdot A_i+\cdots+c_rA_r\cdot A_i=O\cdot A_i=0.$$

Como $A_j\cdot A_i=0$ si $j\neq i$, hallamos que

$$c_iA_i\cdot A_i=0.$$

Pero, por suposición, $A_i\cdot A_i\neq 0$. Por tanto, $c_i=0$, como se quería demostrar.

8. (a)
$$\begin{aligned}\|A+B\|^2+\|A-B\|^2&=(A+B)\cdot(A+B)+(A-B)\cdot(A-B)\\&=A^2+2A\cdot B+B^2+A^2-2A\cdot B+B^2\\&=2A^2+2B^2=2\|A\|^2+2\|B\|^2\end{aligned}$$

9. $\|A-B\|^2=A^2-2A\cdot B+B^2=\|A\|^2-2\|A\|\,\|B\|\cos\theta+\|B\|^2$

I, §5, pág. 33

1. (a) Sea $A=P_2-P_1=(-5,-2,3)$. Una representación paramétrica de la recta es $X(t)=P_1+tA=(1,3,-1)+t(-5,-2,3)$.
(b) $(-1,5,3)+t(-1,-1,4)$

2. $X=(1,1,-1)+t(3,0,-4)$ **3.** $X=(-1,5,2)+t(-4,9,1)$

4. (a) $(-\frac{3}{2},4,\frac{1}{2})$ (b) $(-\frac{2}{3},\frac{11}{3},0)$, $(-\frac{7}{3},\frac{13}{3},1)$ (c) $(0,\frac{17}{5},-\frac{2}{5})$ (d) $(-1,\frac{19}{5},\frac{1}{5})$

5. $P+\dfrac{1}{2}(Q-P)=\dfrac{P+Q}{2}$

I, §6, pág. 39

1. Los vectores normales $(2,3)$ y $(5,-5)$ no son perpendiculares entre sí debido a que su producto interior $10-15=-5$ no es igual a 0.

2. Los vectores normales son $(-m,1)$ y $(-m',1)$ y su producto interior es igual a $mm'+1$. Los vectores son perpendiculares si, y sólo si, este producto interior es igual a 0, lo que equivale a que $mm'=-1$.

3. $y=x+8$ **4.** $4y=5x-7$ **6.** (c) y (d)

7. (a) $x-y+3z=-1$ (b) $3x+2y-4z=2\pi+26$ (c) $x-5z=-33$

8. (a) $2x+y+2z=7$ (b) $7x-8y-9z=-29$ (c) $y+z=1$

9. $(3,-9,-5)$, $(1,5,-7)$ (Otros serían múltiplos constantes de éstos.)

10. $(-2,1,5)$ **11.** $(11,13,-7)$

12. (a) $X=(1,0,-1)+t(-2,1,5)$
(b) $X=(-10,-13,7)+t(11,13,-7)$ o también $(1,0,0)+t(11,13,-7)$

13. (a) $-\frac{1}{3}$ (b) $-\frac{2}{\sqrt{42}}$ (c) $\frac{4}{\sqrt{66}}$ (d) $-\frac{2}{\sqrt{18}}$

14. (a) $(-4,\frac{11}{2},\frac{15}{2})$ (b) $(\frac{25}{13},\frac{10}{13},-\frac{9}{13})$ **15.** $(1,3,-2)$

16. (a) $\frac{8}{\sqrt{35}}$ (b) $\frac{13}{\sqrt{21}}$

II, §1, pág. 45

1. $A+B=\begin{pmatrix}0&7&1\\0&1&1\end{pmatrix}$, $3B=\begin{pmatrix}-3&15&-6\\3&3&-3\end{pmatrix}$

$-2B=\begin{pmatrix}2&-10&4\\-2&-2&2\end{pmatrix}$, $A+2B=\begin{pmatrix}-1&12&-1\\1&2&0\end{pmatrix}$

$2A+B=\begin{pmatrix}1&9&4\\-1&1&3\end{pmatrix}$, $A-B=\begin{pmatrix}2&-3&5\\-2&-1&3\end{pmatrix}$

$A-2B=\begin{pmatrix}3&-8&7\\-3&-2&4\end{pmatrix}$, $B-A=\begin{pmatrix}-2&3&-5\\2&1&-3\end{pmatrix}$

2. $A+B=\begin{pmatrix}0&0\\2&-2\end{pmatrix}$, $3B=\begin{pmatrix}-3&3\\0&-9\end{pmatrix}$, $-2B=\begin{pmatrix}2&-2\\0&6\end{pmatrix}$,

$A+2B=\begin{pmatrix}-1&1\\2&-5\end{pmatrix}$, $A-B=\begin{pmatrix}2&-2\\2&4\end{pmatrix}$, $B-A=\begin{pmatrix}-2&2\\-2&-4\end{pmatrix}$

Renglones de A: $(1,2,3)$, $(-1,0,2)$

Columnas de A: $\begin{pmatrix}1\\-1\end{pmatrix}$, $\begin{pmatrix}2\\0\end{pmatrix}$, $\begin{pmatrix}3\\2\end{pmatrix}$

Renglones de B: $(-1,5,-2)$, $(1,1,-1)$

Columnas de B: $\begin{pmatrix}-1\\1\end{pmatrix}$, $\begin{pmatrix}5\\1\end{pmatrix}$, $\begin{pmatrix}-2\\-1\end{pmatrix}$

Renglones de A: $(1,-1),(2,1)$ Columnas de A: $\begin{pmatrix}1\\2\end{pmatrix},\begin{pmatrix}-1\\1\end{pmatrix}$

Renglones de B: $(-1,1),(0,-3)$ Columnas de B: $\begin{pmatrix}-1\\0\end{pmatrix},\begin{pmatrix}1\\-3\end{pmatrix}$

4. (a) ${}^tA=\begin{pmatrix}1&-1\\2&0\\3&2\end{pmatrix}$, ${}^tB=\begin{pmatrix}-1&1\\5&1\\-2&-1\end{pmatrix}$

(b) ${}^tA=\begin{pmatrix}1&2\\-1&1\end{pmatrix}$, ${}^tB=\begin{pmatrix}-1&0\\1&-3\end{pmatrix}$

5. Sea $c_{ij}=a_{ij}+b_{ij}$. La componente ij de ${}^t(A+B)$ es $c_{ji}=a_{ji}+b_{ji}$, la cual es la suma de la componente ji de A más la componente ji de B.

7. Son iguales 8. $\begin{pmatrix}0&2\\0&-2\end{pmatrix}$, es igual

9. $A+{}^tA=\begin{pmatrix}2&1\\1&2\end{pmatrix}$, $B+{}^tB=\begin{pmatrix}-2&1\\1&-6\end{pmatrix}$

10. (a) ${}^t(A+{}^tA)={}^tA+{}^{tt}A={}^tA+A=A+{}^tA$.
(b) ${}^t(A-{}^tA)={}^tA-{}^{tt}A=-(A-{}^tA)$.
(c) Los elementos diagonales son iguales a 0, ya que satisfacen

$$a_{ii}=-a_{ii}.$$

II, §2, pág. 55

1. $IA=AI=A$ 2. O

3. (a) $\begin{pmatrix}3&2\\4&1\end{pmatrix}$ (b) $\begin{pmatrix}10\\14\end{pmatrix}$ (c) $\begin{pmatrix}33&37\\11&-18\end{pmatrix}$

5. $AB=\begin{pmatrix}4&2\\5&-1\end{pmatrix}$, $BA=\begin{pmatrix}2&4\\4&1\end{pmatrix}$

6. $AC=CA=\begin{pmatrix}7&14\\21&-7\end{pmatrix}$, $BC=CB=\begin{pmatrix}14&0\\7&7\end{pmatrix}$
Si $C=xI$, donde x es un número, entonces $AC=CA=xA$.

7. $(3,1,5)$, primer renglón

8. Segundo renglón, tercer renglón, i-ésimo renglón

9. (a) $\begin{pmatrix}2\\4\end{pmatrix}$ (b) $\begin{pmatrix}4\\6\end{pmatrix}$ (c) $\begin{pmatrix}3\\5\end{pmatrix}$

10. (a) $\begin{pmatrix}3\\1\\2\end{pmatrix}$ (b) $\begin{pmatrix}12\\3\\9\end{pmatrix}$ (c) $\begin{pmatrix}5\\4\\8\end{pmatrix}$

11. Segunda columna de A 12. j-ésima columna de A

13. (a) $\begin{pmatrix} 4 \\ 9 \\ 5 \end{pmatrix}$ (b) $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ (c) $\begin{pmatrix} x_2 \\ 0 \end{pmatrix}$ (d) $\begin{pmatrix} 0 \\ x_1 \end{pmatrix}$

14. (a) $\begin{pmatrix} a & ax+b \\ c & cx+d \end{pmatrix}$. Sume un múltiplo de la primera columna a la segunda. Los otros casos son similares.

16. (a) $A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $A^3 =$ matriz O. Si $B = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ entonces

$$B^2 = \begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{y} \quad B^4 = O.$$

(b) $A^2 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$, $A^3 = \begin{pmatrix} 1 & 3 & 6 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$, $A^4 = \begin{pmatrix} 1 & 4 & 10 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}$

17. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 27 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 81 \end{pmatrix}$

18. Matriz diagonal cuya diagonal es $a_1^k, a_2^k \dots, a_n^k$.

19. $0, 0$

20. (a) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

(b) $\begin{pmatrix} a & b \\ -a^2/b & -a \end{pmatrix}$ para cualesquiera $a, b \neq 0$; si $b = 0$, entonces $a = 0$.

21. (a) La inversa es $I + A$.
(b) Multiplique $I - A$ por $I + A + A^2$ en cada lado. ¿Qué obtiene?

22. (a) Multiplique cada miembro de la relación $B = TAT^{-1}$ por la izquierda por T^{-1} y por la derecha por T. Obtenemos

$$T^{-1}BT = T^{-1}TAT^{-1}T = IAI = A.$$

En consecuencia, existe una matriz, a saber, T^{-1}, tal que $T^{-1}BT = A$. Esto significa que B es semejante a A.

(b) Suponga que A tiene inversa A^{-1}. Entonces $TA^{-1}T^{-1}$ es una inversa de B, ya que

$$TA^{-1}T^{-1}B = TA^{-1}T^{-1}TAT^{-1} = TA^{-1}AT^{-1} = TT^{-1} = I.$$

Y, en forma similar, $BTA^{-1}T^{-1} = I$.

(c) Considere la transpuesta de la relación $B = TAT^{-1}$. Obtenemos

$${}^tB = {}^tT^{-1}\,{}^tA\,{}^tT.$$

Esto significa que ${}^tB$ es semejante a ${}^tA$, debido a que existe una matriz, a saber, ${}^tT^{-1} = C$, tal que ${}^tB = CAC^{-1}$.

23. Los elementos diagonales son $a_{11}b_{11}, \ldots, a_{nn}b_{nn}$. Se multiplican componente a componente.

24. $\begin{pmatrix} 1 & a+b \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & na \\ 0 & 1 \end{pmatrix}$

25. $\begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix}$

26. Multiplique AB a cada lado por $B^{-1}A^{-1}$. ¿Qué obtiene? Observe el orden en el que se consideran las inversas.

27. (a) La fórmula de adición para el coseno es

$$\cos(\theta_1 + \theta_2) = \cos\theta_1 \cos\theta_2 - \operatorname{sen}\theta_1 \operatorname{sen}\theta_2.$$

Esta fórmula y la del seno darán por resultado lo que usted quiere.

(b) $A(\theta)^{-1} = A(-\theta)$. Multiplique $A(\theta)$ por $A(-\theta)$; ¿qué obtiene?

(c) $A^n = \begin{pmatrix} \cos n\theta & -\operatorname{sen} n\theta \\ \operatorname{sen} n\theta & \cos n\theta \end{pmatrix}$. Puede probar esto por inducción. Considere el producto de A^n por A. ¿Qué obtiene?

28. (a) $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ (b) $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ (d) $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$

(e) $\frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}$ (f) $\frac{1}{2}\begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix}$ (g) $\frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix}$

29. $\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}$ **30.** $\frac{1}{\sqrt{2}}(-1, 3)$ **31.** $(-3, -1)$

32. Las coordenadas de Y están dadas por

$$y_1 = x_1 \cos\theta - x_2 \operatorname{sen}\theta,$$

$$y_2 = x_1 \operatorname{sen}\theta + x_2 \cos\theta.$$

Halle $y_1^2 + y_2^2$ mediante un desarrollo, usando simple aritmética. Muchos términos se cancelarán para dejar $x_1^2 + x_2^2$.

33. (a) $\begin{pmatrix} 1 & 4 & 2 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 2 & 3 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

34. (a) Intercambie el primer renglón de A con el segundo.
(b) Intercambie el segundo renglón de A con el tercero.
(c) Sume cinco veces el segundo renglón de A al cuarto.
(d) Sume -2 veces el segundo renglón de A al tercero.

35. (a) Multiplique el primer renglón de A por 3.
(b) Sume 3 veces el tercer renglón al primero.
(c) Reste 2 veces el primer renglón del segundo.
(d) Reste 2 veces el segundo renglón del tercero.

36. (a) Ponga el renglón s de A en el lugar r, ceros en todo lo demás.
(b) Intercambie los renglones r y s, ponga ceros en todo lo demás.
(c) Intercambie los renglones r y s.

37. (a) Sume 3 veces el renglón s al renglón r.
(b) Sume c veces el renglón s al renglón r.

II, §3, pág. 65

1. Sea $X = (x_1, \ldots, x_n)$. Entonces $X \cdot E_i = x_i$, de manera que si éste es 0 para todo i, entonces $x_i = 0$

3. $X \cdot (c_1A_1 + \cdots + c_nA_n) = c_1X \cdot A_1 + \cdots + c_nX \cdot A_n = 0.$

II, §4, pág. 72

(Hay varias respuestas posibles para la forma escalonada por renglones; mostramos una de ellas, pero otras también son correctas.)

1. (a) $\begin{pmatrix} 1 & 2 & -5 \\ 0 & 9 & -26 \\ 0 & 0 & 0 \end{pmatrix}$ y también $\begin{pmatrix} 1 & 0 & \frac{7}{9} \\ 0 & 1 & -\frac{26}{9} \\ 0 & 0 & 0 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{pmatrix}$ y también $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

2. (a) $\begin{pmatrix} 1 & -2 & 3 & -1 \\ 0 & 3 & -4 & 4 \\ 0 & 0 & 7 & -10 \end{pmatrix}$ y también $\begin{pmatrix} 1 & 0 & 0 & \frac{15}{7} \\ 0 & 1 & 0 & -\frac{4}{7} \\ 0 & 0 & 1 & -\frac{10}{7} \end{pmatrix}$

(b) $\begin{pmatrix} 2 & 0 & -7 & 5 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ y también $\begin{pmatrix} 1 & 0 & -\frac{7}{2} & \frac{5}{2} \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

3. (a) $\begin{pmatrix} 1 & 2 & -1 & 2 & 1 \\ 0 & 0 & 3 & -6 & 1 \\ 0 & 0 & 0 & -6 & 1 \end{pmatrix}$ o también $\begin{pmatrix} 1 & 2 & 0 & 0 & \frac{4}{3} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{6} \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 3 & -1 & 2 \\ 0 & 11 & -5 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ o también $\begin{pmatrix} 1 & 0 & \frac{4}{11} & \frac{13}{11} \\ 0 & 1 & -\frac{5}{11} & \frac{3}{11} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

II, §5, pág. 79

1. (a) $-\frac{1}{20}\begin{pmatrix} 4 & 1 & -7 \\ -4 & -6 & 2 \\ -12 & 2 & 6 \end{pmatrix}$ (b) $\frac{1}{5}\begin{pmatrix} 2 & 23 & -11 \\ 1 & 19 & -8 \\ 0 & -15 & 5 \end{pmatrix}$

(c) $\frac{1}{4}\begin{pmatrix} 3 & 2 & -9 \\ 1 & 2 & -3 \\ -2 & -4 & 10 \end{pmatrix}$ (d) $\frac{1}{5}\begin{pmatrix} 5 & -16 & 3 \\ 0 & 7 & -1 \\ 0 & -2 & 1 \end{pmatrix}$

(e) $-\frac{1}{76}\begin{pmatrix} 0 & -19 & 0 \\ -32 & -14 & 12 \\ 28 & 17 & -20 \end{pmatrix}$

(f) $-\frac{1}{14}\begin{pmatrix} -17 & 7 & -9 \\ 11 & -7 & 5 \\ 13 & -7 & 11 \end{pmatrix}$

2. El efecto de multiplicar por I_{rs} consiste en poner al renglón s en el lugar r y ceros en todo lo demás. Así, el renglón s de $I_{rs}A$ es 0. Al multiplicar por I_{rs} una vez más se pone 0 en el lugar r. y 0 en todo lo demás, de manera que $I_{rs}^2 = O$.

3. Tenemos $E_{rs}(c) = I + cI_{rs}$ y $E_{rs}(c') = I + c'I_{rs}$, de manera que

$$\begin{aligned} E_{rs}(c)E_{rs}(c') &= (I + cI_{rs})(I + c'I_{rs}) \\ &= I + cI_{rs} + c'I_{rs} + cc'I_{rs}^2 \\ &= I + (c + c')I_{rs} \qquad \text{porque} \qquad I_{rs}^2 = O \\ &= E_{rs}(c + c'). \end{aligned}$$

III, §1, pág. 87

1. Sean B y C perpendiculares a A_i para todo i. Entonces

$$(B + C) \cdot A_i = B \cdot A_i + C \cdot A_i = 0 \qquad \text{para todo } i.$$

Además, para cualquier número x,

$$(xB) \cdot A_i = x(B \cdot A_i) = 0.$$

Por último, $O \cdot A_i = 0$ para todo i. Esto prueba que W es un subespacio.

2. (c) Sea W el conjunto de todas las (x, y) tales que $x + 4y = 0$. Entonces, los elementos de W son de la forma $(-4y, y)$. Al poner $y = 0$ se muestra que $(0, 0)$ está en W. Si $(-4y, y)$ y $(-4y, y')$ están en W, entonces su suma es $(-4(y+y'), y+y')$, y, por tanto, está en W. Si c es un número, entonces $c(-4y, y) = (-4cy, cy)$, que pertenece a W. Por tanto, W es un subespacio.

4. Suponga que v_1 y v_2 están en la intersección $U \cap W$. Entonces su suma $v_1 + v_2$ está tanto en U (ya que v_1 y v_2 están en U) como en W (porque v_1 y v_2 están en W) de manera que está en la intersección $U \cap W$. Dejamos al lector las otras condiciones.

Ahora probemos parcialmente que $U + W$ es un subespacio. Sean u_1 y u_2 elementos de U y w_1 y w_2 elementos de W. Entonces

$$(u_1 + w_1) + (u_2 + w_2) = u_1 + u_2 + w_1 + w_2,$$

y ésta tiene la forma $u + w$, con $u = u_1 + u_2$ en U y $w = w_1 + w_z$ en W. Por tanto, la suma de dos elementos de $U + W$ también está en $U + W$. Dejamos al lector las otras condiciones.

Con respecto a (iii), dado un vector arbitrario (s, t), resuelva por eliminación el sistema de ecuaciones lineales que proviene de

$$xA + yC = (s, t).$$

Al hacerlo el lector hallará que es necesario que $ad - bc \neq 0$.

6. Véase el Capítulo I, sección §4, Ejercicio 7.

9. $(3,5)$ **10.** $(-5,3)$

III, §4, pág. 102

2. (a) $A-B, (1,-1)$ (b) $\frac{1}{2}A+\frac{3}{2}B, (\frac{1}{2},\frac{3}{2})$
(c) $A+B, (1,1)$ (d) $3A+2B, (3,2)$

3. (a) $(\frac{1}{3},-\frac{1}{3},\frac{1}{3})$ (b) $(1,0,1)$ (c) $(\frac{1}{3},-\frac{1}{3},-\frac{2}{3})$

4. Suponga que $ad-bc\neq 0$. Sean $A=(a,b)$ y $C=(c,d)$. Suponga que tenemos

$$xA+yC=O.$$

Esto significa, en términos de coordenadas, que

$$xa+yc=0,$$

$$xb+yd=0.$$

Multiplique la primera ecuación por d, la segunda por c y reste. Hallamos que

$$x(ad-bc)=0.$$

Como $ad-bc\neq 0$, esto implica que $x=0$. Una eliminación similar muestra que $y=0$. Esto prueba (i).

Recíprocamente, suponga que A y C son linealmente independientes. Entonces ninguna de ellas puede ser $(0,0)$ (de otra manera se escoge $x,y\neq 0$ y se obtiene $xA+yC=0$, lo cual es imposible). Digamos que b o $d\neq 0$. Entonces

$$d(a,b)-b(c,d)=(ad-bc,0).$$

Como se supone que A y C son linealmente independientes, el miembro de la derecha no puede ser 0, por lo que $ad-cb\neq 0$. El argumento es similar si a o $c\neq 0$.

Con respecto a (iii), dado un vector arbitrario (s,t), resuelva el sistema de ecuaciones lineales que proviene de $xA+yC=(s,t)$, por eliminación. El lector encontrará precisamente que necesita que $ad-bc\neq 0$ para hacerlo así.

6. Vea el Capítulo I, sección §4, Ejercicio 7.

11. Base posible: $\begin{pmatrix}1&0\\0&0\end{pmatrix}, \begin{pmatrix}0&1\\0&0\end{pmatrix}, \begin{pmatrix}0&0\\1&0\end{pmatrix}, \begin{pmatrix}0&0\\0&1\end{pmatrix}$

12. $\{E_{ij}\}$ donde E_{ij} tiene componente 1 en el lugar (i,j) y 0 en todos los demás. Estos elementos generan $\mathrm{Mat}(m\times n)$, debido a que, dada cualquier matriz $A=(a_{ij})$, la podemos escribir como una combinación lineal

$$A=\textstyle\sum_i\sum_j a_{ij}E_{ij}.$$

Además, si

$$O=\textstyle\sum_i\sum_j a_{ij}E_{ij},$$

entonces debemos tener que $a_{ij}=0$ para todos los índices i y j, por lo que los elementos E_{ij} son linealmente independientes.

13. E_i, donde E_i es la matriz de $n \times n$ cuyo término ii es 1 y todos los demás son 0.

14. Se puede escoger una base que conste de los elementos E_{ij} que tengan componente ij igual a 1 para $i \leq j$ y todas las demás componentes iguales a 0. El número de tales elementos es

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

15. (a) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

(b) $\left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \right\}$

16. Se puede tomar como base para el espacio $\text{Sim}(n \times n)$ de las matrices simétricas de $n \times n$ a los elementos E_{ij}, donde $i \leq j$, que tienen componente ij igual a 1, componente ji igual a 1 y componente rs igual a 0 si $(r, s) \neq (i, j)$ o (j, i). La prueba de que éstos generan a $\text{Sim}(n \times n)$ y que son linealmente independientes es semejante a la prueba que aparece en el Ejercicio 12.

III, §5, pág. 107

1. (a) 4 (b) mn (c) n (d) $n(n+1)/2$ (e) 3 (f) 6 (g) $n(n+1)/2$

2. 0, 1 ó 2, por el Teorema 5.8. El subespacio consta sólo del O si, y sólo si, tiene dimensión 0. Si el subespacio tiene dimensión 1, sea v_1 una base. Entonces el subespacio consta de todos los elementos tv_1, para todos los números t, por lo que, por definición, es una recta. Si el subespacio tiene dimensión igual a 2, sean v_1 y v_2 una base. Entonces el subespacio consta de todos los elementos $t_1 v_1 + t_2 v_2$, donde t_1 y t_2 son números, por lo que es, por definición, un plano.

3. 0, 1, 2 ó 3 por el Teorema 5.8.

III, §6, pág. 114

1. (a) 2 (b) 2 (c) 2 (d) 1 (e) 2 (f) 3 (g) 3 (h) 2 (i) 2

IV, §1, pág. 118

1. (a) $\cos x$ (b) e^x (c) $1/x$ **2.** $(1/\sqrt{2}, 1/\sqrt{2})$

3. (a) 11 (b) 13 (c) 6

4. (a) $(e, 1)$ (b) $(1, 0)$ (c) $(1/e, -1)$

5. (a) $(e+1, 3)$ (b) $(e^2+2, 6)$ (c) $(1, 0)$

6. (a) $(2, 0)$ (b) $(\pi e, \pi)$

7. (a) 1 (b) 11

8. Elipse $9x^2 + 4y^2 = 36$ **9.** Recta $x = 2y$

10. Círculo $x^2 + y^2 = e^2$, círculo $x^2 + y^2 = e^{2c}$

11. Cilindro, radio 1, eje $z =$ eje del cilindro 12. Círculo $x^2 + y^2 = 1$.

IV, §2, pág. 125

1. Todos excepto (c) y (g)

2. Sólo el Ejercicio 8.

5. Como $AX = BX$ para todo X, esta relación es verdadera, en particular cuando $X = E^j$ es el j-ésimo vector unitario. Pero entonces $AE^j = A^j$ es la columna j de A y $BE^j = B^j$ es la columna j de B, por lo que $A^j = B^j$ para todo j. Esto prueba que $A = B$.

6. Sólo $u = O$, ya que $T_u(O) = u$ y, si T_u es lineal, entonces debemos tener $T_u(O) = O$.

7. La recta S se puede representar en la forma $P + tv_1$ con todos los números t. Luego, $L(S)$ consta de todos los puntos

$$L(P) + tL(v_1).$$

Si $L(v_1) = O$, éste es simplemente un punto sencillo. Si $L(v_1) \neq O$, ésta es una recta. Otros casos se resuelven en forma similar.

8. Es un paralelogramo cuyos vértices son B, $3A$, $3A + B$, O.

9. Es un paralelogramo cuyos vértices son 0, $2B$, $5A$, $5A + 2B$.

10. (a) $(-1, -1)$ (b) $(-2/3, 1)$ (c) $(-2, -1)$

11. (a) $(4, 5)$ (b) $(11/3, -3)$ (c) $(4, 2)$

12. Suponga que tenemos una relación $\sum x_i v_i = O$. Aplique F; obtenemos $\sum x_i F(v_i) = \sum x_i w_i = O$. Como los w_i son linealmente independientes, se infiere que todo $x_i = 0$.

13. (a) Sea v un elemento arbitrario de V. Como $F(v_0) \neq 0$, existe un número c tal que

$$F(v) = cF(v_0),$$

a saber, $c = F(v)/F(v_0)$. Entonces $F(v - cv_0) = 0$, por tanto, haga $w = v - cv_0$. Hemos escrito $v = w + cv_0$, tal como se deseaba.

(b) W es un subespacio, por el Ejercicio 3. Por la parte (a), los elementos v_0, $v_1, \ldots, v_n$ generan a V. Suponga que existe una relación lineal

$$c_0 v_0 + c_1 v_1 + \cdots + c_n v_n = O.$$

Aplique F. Obtenemos $c_0 F(v_0) = 0$. Como $F(v_0) \neq 0$, se infiere que $c_0 = 0$. Pero entonces $c_i = 0$ para $i = 1, \ldots, n$, debido a que $v_1, \ldots, v_n$ forman una base de W.

IV, §3, pág. 131

1 y 2. Si U es un espacio de V, entonces $\dim L(U) \leq \dim U$. Por tanto, la imagen de un subespacio bidimensional es 0, 1 ó 2. Una recta o plano es de la forma $P + U$, donde U tiene dimensión 1 ó 2. Su imagen es de la forma $L(P) + L(U)$, de manera que ahora los asertos están claros.

3. (a) Por la fórmula de la dimensión, la imagen de F tiene dimensión n. Por el Teorema 4.6 del Capítulo III, la imagen debe ser todo W.
(b) Es similar.

4. Use la fórmula de la dimensión.

5. Como $L(v_0 + u) = L(v_0)$ si u está en $\operatorname{Ker} L$, todo elemento de la forma $v_0 + u$ es una solución. Recíprocamente, sea v una solución de $L(v) = w$. Entonces

$$L(v - v_0) = L(v) - L(v_0) = w - w = O,$$

por lo que $v - v_0 = u$ está en el núcleo y $v = v_0 + u$.

6. Funciones constantes.

7. $\operatorname{Ker} D^2$ = polinomios de grad ≤ 1, $\operatorname{Ker} D^n$ = polinomios de grad $\leq n - 1$.

8. (a) Múltiplos constantes de e^x (b) Múltiplos constantes de e^{ax}.

9. (a) $n - 1$ (b) $n^2 - 1$.

10. $A = \frac{A + A'}{2} + \frac{A - {}^tA}{2}$. Si $A = B + C = B_1 + C_1$, entonces

$$B - B_1 = C_1 - C.$$

Pero $B - B_1 + C_1 - C$ es simétrica y antisimétrica, y por tanto O, ya que cada componente es igual a su propio negativo.

11. (c) Tomar la transpuesta de $(A + {}^tA)/2$ demuestra que ésta es una matriz simétrica. Recíprocamente, dada una matriz simétrica B, vemos que $B = P(B)$, de manera que B está en la imagen de P.
(d) $n(n-1)/2$.
(e) Una base para las matrices antisimétricas consta de las matrices E_{ij} con $i < j$, que tiene componente ij igual a 1, componente ji igual a -1 y todas las otras componentes iguales a 0.

12. Semejante a 11.

13 y **14.** Similar a 11 y 12.

15. (a) 0 (b) $m + n$, $\{(u_i, 0), (0, w_j)\}$; $i = 1, \ldots, m$; $j = 1, \ldots, n$. Si $\{u_i\}$ es una base de U y $\{w_j\}$ es una base de W.

16. (b) Se ve claramente que la imagen está contenida en $U + W$. Dado un elemento arbitrario $u + w$ con u en U y w en W, podemos escribirlo en la forma $u + w = u - (-w)$, lo que demuestra que está en la imagen de L.
(c) El núcleo de L consta de aquellos elementos (u, w) tales que $u - w = O$, de manera que $u = w$. En otras palabras, consta de las parejas (u, u) y u debe pertenecer a U y a W, por tanto, está en su intersección. Si $\{u_1, \ldots, u_r\}$ es una base para $U \cap W$, entonces $\{(u_1, u_1), \ldots, (u_r, u_r)\}$ es una base del núcleo de L. La dimensión es la misma que la dimensión de $U \cap W$. Después se aplica la fórmula de la dimensión que aparece en el texto.

IV, §4, pág. 139

1. $n - 1$ **2.** 4 **3.** $n - 1$

4. (a) dim. $= 1$ base $= (1, -1, 1)$
(b) dim. $= 2$ base $= (1, 1, 0)(0, 1, 1)$
(c) dim. $= 1$ base $= \left(\frac{\pi - 3}{10}, \frac{\pi + 2}{5}, 1\right)$
(d) dim. $= 0$

5. (a) 1 (b) 1 (c) 0 (d) 2

6. Un teorema establece que

$$\dim V = \dim \operatorname{Im} L + \dim \operatorname{Ker} L.$$

Como $\dim \operatorname{Ker} L \geq 0$, se infiere la desigualdad deseada.

7. Una prueba (hay otras): rango $A = \dim \operatorname{Im} L_A$. Pero $L_{AB} = L_A \circ L_B$. Por tanto, la imagen de L_{AB} está contenida en la imagen de L_A. En consecuencia, rango $AB \leq$ rango A.

Con respecto a la segunda desigualdad, observe que el rango de una matriz es igual al rango de su transpuesta, debido a que el rango por columnas es igual al rango por renglones. Por ese motivo,

$$\text{rango } AB = \text{rango } {}^tB{}^tA.$$

Ahora aplique la primera desigualdad para obtener rango ${}^tB{}^tA \leq$ rango ${}^tB =$ rango B.

IV, §5, pág. 145

1. (a) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

(b) $3I$ (d) 71 (e) $-I$ (f) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

2. cI, donde I es la matriz unitaria de $n \times n$.

3. (a) $\begin{pmatrix} 1 & -4 & 3 \\ -3 & 2 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 3 & -2 & 1 \\ 4 & -1 & 5 \end{pmatrix}$

4. (a) $\begin{pmatrix} 0 & 1 & -2 \\ 3 & -2 & 4 \\ -1 & 1 & 5 \end{pmatrix}$ (b) $\begin{pmatrix} 3 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & 5 \end{pmatrix}$
(c) $\begin{pmatrix} -2 & 0 & 1 \\ 0 & 0 & 0 \\ 7 & -1 & 0 \end{pmatrix}$

5. Sea $Lv_i = \sum c_{ij} w_j$. Sea $C = (c_{ij})$. La matriz asociada es ${}^tC$ y el efecto de L sobre un vector X de coordenadas es ${}^tCX$.

V, §1, pág. 151

1. Sea $C = A - B$. Entonces $CX = O$ para todo X. Considere $X = E^j$ como el j-ésimo vector unitario para $j = 1, \ldots, n$. Entonces $CE^j = C^j$ es la j-ésima columna de C. Por suposición, $CE^j = O$ para todo j, de manera que $C = O$.

2. Use la distributividad y el hecho de que $F \circ L = L \circ F$.

3. La prueba es igual a la que se hace con números.

4. $P^2 = \frac{1}{4}(I+T)^2 = \frac{1}{4}(I^2 + 2TI + T^2) = \frac{1}{4}(2I + 2T) = \frac{1}{2}(I+T) = P$. La parte (b) se deja al lector. Con respecto a la parte (c), vea el siguiente problema.

5 (a) $Q^2 = (I-P)^2 = I^2 - 2IP + P^2 = I - 2P + P = I - P = Q$.
(b) Sea $v \in \operatorname{Im} P$ tal que $v = Pw$ para algún w. Entonces $Qv = QP_v = 0$, ya que $QP = (I-P)P = P - P^2 = P - P = O$. Por tanto, $\operatorname{Im} P \subset \operatorname{Ker} Q$. Recíprocamente, sea $v \in \operatorname{Ker} Q$ tal que $Qv = 0$. Entonces $(I-P)v = 0$, por lo que $v - Pv = 0$, y $v = Pv$, por lo que $v \in \operatorname{Im} P$; en consecuencia, $\operatorname{Ker} Q \subset \operatorname{Im} P$.

6. Sea $v \in V$. Entonces $v = v - Pv + Pv$, y $v - Pv \in \operatorname{Ker} P$, ya que

$$P(v - Pv) = Pv - P^2 v = Pv - Pv = 0.$$

Además $Pv \in \operatorname{Im} P$, lo que prueba (a).
Con respecto a (b), sea $v \in \operatorname{Im} P \cap \operatorname{Ker} P$. Como $v \in \operatorname{Im} P$, existe $w \in V$ tal que $v = Pw$. Como $v \in \operatorname{Ker} P$, obtenemos

$$0 = Pv = P^2 w = Pw = v,$$

de modo que $v = 0$, con lo que también se ha probado (b).

7. Suponga que $u + w = u_1 + w_1$. Entonces $u - u_1 = w_1 - w$. Pero $u - u_1 \in U$ y $w_1 - w \in W$, debido a que U y W son subespacios. Por suposición, $U \cap W = \{0\}$, y concluimos que $u - u_1 = 0 = w_1 - w$, por lo que $u = u_1$ y $w = w_1$.

8. $P^2(u, w) = P(u, 0) = (u, 0) = P(u, w)$. De modo que $P^2 = P$.

9. La dimensión de un subespacio es $\leq$ la dimensión del espacio. Entonces

$$\operatorname{Im} F \circ L \leq \operatorname{Im} F, \quad \text{de modo que} \quad \dim \operatorname{Im} F \circ L \leq \dim \operatorname{Im} F$$

y así, rango $F \circ L \leq$ rango F. Esto prueba una de las fórmulas. Con respecto a la otra, considere F como una combinación lineal definida sobre $\operatorname{Im} L$. Entonces

$$\text{rango } F \circ L = \dim \operatorname{Im} F \circ L \leq \dim \operatorname{Im} L = \text{rango } L.$$

V, §2, pág. 156

1. $R_\theta^{-1} = R_{-\theta}$ debido a que $R_\theta \circ R_{-\theta} = R_{\theta - \theta} = R_0 = I$. La matriz asociada con R_θ^{-1} es

$$\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

debido a que $\cos(-\theta) = \cos\theta$.

3. Las composiciones siguientes son la identidad:

$$F \circ G \circ G^{-1} \circ F^{-1} = I \qquad \text{y} \qquad G^{-1} \circ F^{-1} \circ F \circ G = I.$$

4, **5**, **6**. En cada caso demuestre que el núcleo es igual a O y aplique el teorema apropiado.

7 al **10**. La prueba es similar a la misma prueba para las matrices, usando la distributividad. En **7** tenemos

$$(I - L) \circ (I + L) = I^2 - L^2 = I.$$

Con respecto a **8**, tenemos que $L^2 + 2L = -I$, de modo que $L(-L-2I)L = I$, por lo que $L^{-1} = -L - 2I$.

11. Es suficiente probar que v y w son linealmente independientes. Suponga que $xv + yw = O$. Aplique L. Entonces

$$L(w) = L\big(L(v)\big) = O$$

ya que $L^2 = O$. En consecuencia, $L(xv) = xL(v) = O$. Como $L(v) \neq O$, se infiere que $x = 0$. Entonces $y = 0$ por que $w \neq O$.

12. F es inyectiva, su núcleo es 0.
   (b) F no es suprayectiva; por ejemplo $(1, 0, 0, \ldots)$ no está en la imagen.
   (c) Sea $G(x_1, x_2, \ldots) = (x_2, x_3, \ldots)$ (elimina la primera coordenada).
   (d) No; en caso contrario F tendría una inversa, lo cual no sucede.

13. La linealidad se comprueba fácilmente. Para demostrar que L es inyectiva, es suficiente mostrar que $\operatorname{Ker} L = \{0\}$. Suponga que $L(u, w) = 0$; entonces $u + w = 0$, de manera que $u = -w$. Por suposición, $U \cap W = \{0\}$ y $u \in U$, $-w \in W$, por lo que $u = w = 0$. En consecuencia, $\operatorname{Ker} L = \{0\}$.
   L es suprayectiva debido a que, por suposición, todo elemento de V se puede escribir como la suma de un elemento de U más un elemento de W.

VI, §1, pág. 164

2. Sea $X = {}^t(x, y)$. Entonces

$$\langle X, X\rangle = ax^2 + 2bxy + dy^2 = a\left(x + \frac{b}{a}y\right)^2 + \frac{(ad - b^2)}{a}y^2.$$

Si $ad - b^2 > 0$, entonces $\langle X, X\rangle$ es una suma de cuadrados con coeficientes positivos y uno de los términos no es igual a 0, por lo que $\langle X, X\rangle > 0$. Si $ad - b^2 \leq 0$, entonces dé a y cualquier valor no nulo y sea $x = -by/a$. Entonces $\langle X, X\rangle \leq 0$.
   Si $a = 0$, sea $y = 0$ y dé a x cualquier valor. Entonces ${}^tXAX = 0$.

3. (a) sí (b) no (c) sí (d) no (e) sí (f) sí

4. (a) 2 (b) 4 (c) 8

5. Los elementos diagonales no cambian bajo la transposición, de manera que las trazas de A y ${}^tA$ son las mismas.

6. Sea $AA = (c_{ij})$. Entonces

$$c_{ii} = \sum_{k=1}^{n} a_{ik}a_{ki}.$$

Por tanto,

$$\operatorname{tr}(AA) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik}a_{ki},$$

y $a_{ik}a_{ki} = a_{ik}^2$ debido a que A es simétrica, por lo que la traza es una suma de cuadrados, por tanto, ≥ 0.

8. (a)
$$\operatorname{tr}(AB) = \sum_i \sum_j a_{ij}b_{ji} = \sum_j \sum_i b_{ji}a_{ij}.$$

Pero cualquier par de letras se puede usar para denotar índices en esta suma, la cual se puede escribir en forma más neutral como sigue:

$$\sum_{r=1}^{n} \sum_{s=1}^{n} b_{rs}a_{sr}.$$

Ésta es precisamente la traza $\operatorname{tr}(BA)$.

(b) $\operatorname{tr}(C^{-1}AC) = \operatorname{tr}(ACC^{-1})$, por la parte (a), de modo que es $= \operatorname{tr}(A)$.

VI, §2, pág. 174

1. (a) $\frac{1}{\sqrt{3}}(1,1,-1)$ y $\frac{1}{\sqrt{2}}(1,0,1)$

(b) $\frac{1}{\sqrt{6}}(2,1,1)$, $\frac{1}{5\sqrt{3}}(-1,7,-5)$

2. $\frac{1}{\sqrt{6}}(1,2,1,0)$ y $\frac{1}{\sqrt{39}}(-1,-2,5,3)$

3. $\frac{1}{\sqrt{2}}(1,1,0,0)$, $\frac{1}{2}(1,-1,1,1)$, $\frac{1}{\sqrt{18}}(-2,2,3,1)$

5. $\sqrt{80}(t^2 - 3t/4)$, $\sqrt{3}t$

6. $\sqrt{80}(t^2 - 3t/4)$, $\sqrt{3}t$, $10t^2 - 12t + 3$

9. Use las fórmulas de dimensión. La traza es una aplicación lineal de V en $\mathbf{R}$. Como

$$\dim V = \dim \operatorname{Ker} \operatorname{tr} + \dim \operatorname{Im} \operatorname{tr},$$

se infiere que $\dim W = \dim V - 1$, por lo que $\dim W^{\perp} = 1$ por el Teorema 2.3. Sea I la matriz unitaria de $n \times n$. Entonces $\operatorname{tr}(A) = \operatorname{tr}(AI)$ para toda $A \in V$, de manera que $\operatorname{tr}(AI) = 0$ para $A \in W$. En consecuencia, $I \in W^{\perp}$. Como $W^{\perp}$ tiene dimensión 1, se infiere que I es una base de $W^{\perp}$. (¿Sencillo?)

10. Tenemos que $\langle X, AY\rangle = \langle X, bY\rangle = b\langle X, Y\rangle$, y además

$$\langle X, AY\rangle = \langle AX, Y\rangle = \langle aX, Y\rangle = a\langle X, Y\rangle.$$

Por tanto, $a\langle X, Y\rangle = b\langle X, Y\rangle$. Como $a \neq b$, se infiere que $\langle X, Y\rangle = 0$.

VI, §3, pág. 178

1. Con respecto a todos los vectores columna X y Y, observamos que ${}^tXAY$ es una matriz de 1×1, de modo que es igual a su propia transpuesta. Por consiguiente,

$$\varphi_A(X,Y) = {}^tXAY = {}^t({}^tXAY) = {}^tY{}^tA{}^{tt}X = {}^tYAX = \varphi_A(Y,X).$$

2. Recíprocamente, si $\varphi_A(X,Y) = \varphi_A(Y,X)$, entonces un argumento semejante al anterior muestra que

$${}^tXAY = {}^tX{}^tAY$$

para todos los X, Y. Entonces $A = {}^tA$, por la prueba de unicidad que aparece en el Teorema 3.1.

3. (a) $2x_1y_1 - 3x_1y_2 + 4x_2y_1 + x_2y_2$
   (b) $4x_1y_1 + x_1y_2 - 2x_2y_1 + 5x_2y_2$
   (c) $5x_1y_1 + 2x_1y_2 + \pi x_2y_1 + 7x_2y_2$
   (d) $x_1y_1 + 2x_1y_2 - x_1y_3 - 3x_2y_1 + x_2y_2 + 4x_2y_3 + 2x_3y_1 + 5x_3y_2 - x_3y_3$
   (e) $-4x_1y_1 + 2x_1y_2 + x_1y_3 + 3x_2y_1 + x_2y_2 + x_2y_3 + \quad +2x_3y_1 + 5x_3y_2 + 7x_3y_3$
   (f) $-\frac{1}{2}x_1y_1 + 2x_1y_2 - 5x_1y_3 + x_2y_1 + \frac{2}{3}x_2y_2 + 4x_2y_3 - x_3y_1 + 3x_3y_3$

VII, §2, pág. 190

1. (a) -20 (b) 5 (c) 4 (d) 5 (e) -76 (f) -14

2. (a) -18 (b) 45 (c) 0 (d) 0 (e) 4 (f) 14 (g) 108 (h) 135 (i) 10

3. $a_{11}a_{22}\cdots a_{nn}$ 4. 1

5. Aun en el caso de 3×3, siga las directrices generales. Reste la segunda columna multiplicada por x_1 de la tercera. Reste la primera columna multiplicada por x_1 de la segunda. Obtendrá

$$\begin{vmatrix} 1 & 0 & 0 \\ 1 & x_2 - x_1 & x_2^2 - x_2x_1 \\ 1 & x_3 - x_1 & x_3^2 - x_3x_1 \end{vmatrix}.$$

Al desarrollar conforme al primer renglón se obtiene

$$\begin{vmatrix} x_2 - x_1 & x_2(x_2 - x_1) \\ x_3 - x_1 & x_3(x_3 - x_1) \end{vmatrix}.$$

El lector puede factorizar $x_2 - x_1$ en el primer renglón y $x_3 - x_1$ en el segundo renglón para obtener

$$(x_2 - x_1)(x_3 - x_1)\begin{vmatrix} 1 & x_2 \\ 1 & x_3 \end{vmatrix}.$$

Use después la fórmula para los determinantes de 2×2 a fin de obtener la respuesta deseada.

Ahora haga con detalle el caso de 4×4 para entender el patrón, y luego efectúe el caso de $n \times n$, por inducción.

6. (a) 3 (b) -24 (c) 16 (d) 14 (e) 0 (f) 8
   (g) 40 (h) -10 (i) $\prod_{i=1}^{N} a_{ii}$

7. 1 **8.** $t^2 + 8t + 5$

11. $D(cA) = D(cA^1, cA^2, cA^3) = cD(A^1, cA^2, cA^3) = c^3 D(A^1, A^2, A^3)$ al usar la linealidad con respecto a cada columna.

12. $D(A) = D(cA^1, \ldots, cA^n) = c^n D(A^1, \ldots, A^n)$ usando de nuevo la linealidad con respecto a cada columna.

VII, §3, pág. 195

1. 2 **2.** 2 **3.** 2 **4.** 3 **5.** 4 **6.** 3 **7.** 2 **8.** 3

VII, §4, pág. 198

1. (a) $x = -\frac{1}{3}, y = \frac{2}{3}, z = -\frac{1}{3}$ (b) $x = \frac{5}{12}, y = -\frac{1}{12}, z = \frac{1}{12}$
(c) $x = -\frac{5}{24}, y = \frac{97}{48}, z = \frac{1}{3}, w = -\frac{25}{48}$
(d) $x = \frac{11}{2}, y = \frac{38}{5}, z = \frac{1}{10}, w = 2$

VII, §5, pág. 202

1. Capítulo II, sección §5

2. $\dfrac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

VII, §6, pág. 212

2. (a) 14 (b) 1

3. (a) 11 (b) 38 (c) 8 (d) 1

4. (a) 10 (b) 22 (c) 11 (d) 0

VIII, §1, pág. 217

1. Sea $\begin{pmatrix} x \\ y \end{pmatrix}$ un vector propio. Entonces la multiplicación de matrices muestra que debemos tener $x + ay = \lambda x$ y $y = \lambda y$. Si $y \neq 0$, entonces $\lambda = 1$. Como $a \neq 0$, esto contradice la primera ecuación, por lo cual $y = 0$. Entonces E^1, que es un vector propio, forma una base para el espacio de los vectores propios.

2. Si $A = cI$ es un múltiplo escalar de la identidad, entonces todo el espacio consta de vectores propios. Cualquier base de todo el espacio satisface los requisitos. El único valor propio es el c mismo, para vectores no nulos.

3. El vector unitario E^i es un vector propio, cuyo valor propio es a_{ii}. El conjunto de estos vectores unitarios es una base para todo el espacio.

4. Sea $y = 1$. Entonces, usando multiplicación de matrices, el lector encontrará que ${}^t(x, 1)$ es un vector propio, con valor propio igual a 1. Si $\theta = 0$, los vectores unitarios E^1 y E^2 son vectores propios, con valores propios 1 y -1, respectivamente.

5. Sea $v_2 = {}^t(-1, x)$, de manera que v_2 es perpendicular a v_1 en el Ejercicio 4. La multiplicación de matrices muestra que $Av_2 = -v_2$.

6. Sea $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. Entonces el polinomio característico es $(t-a)^2 + b^2$ y, para que sea igual a 0, debemos tener que $b = 0$ y $t = a$. Pero $a^2 + b^2 = 1$, por lo que $a = \pm 1$.

7. $A(Bv) = ABv = BAv = B\lambda v = \lambda Bv$.

VIII, §2, pág. 227

1. (a) $(t - a_{11}) \dots (t - a_{nn})$ (b) $a_{11}, \dots, a_{nn}$

2. Igual que en el Ejercicio 1.

3. (a) $(t-4)(t+1)$; valores propios $4, -1$; los correspondientes vectores propios $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ y $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, o múltiplos escalares no nulos.
   (b) $(t-1)(t-2)$; valores propios 1, 2; vectores propios $(1, -1/\lambda)$, donde $\lambda = 1$ y $\lambda = 2$.
   (c) $t^2 + 3$; valores propios $\pm\sqrt{-3}$; vectores propios $\big(1, 1/(\lambda - 2)\big)$, donde $\lambda = \sqrt{-3}$ y $\lambda = -\sqrt{-3}$.
   (d) $(t-5)(t+1)$; valores propios $5, -1$; vectores propios $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ y $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$, respectivamente, o múltiplos escalares no nulos.

4. (a) $(t-1)(t-2)(t-3)$; valores propios 1, 2, 3; vectores propios

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -1/2 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

   respectivamente, o múltiplos escalares no nulos.
   (b) $(t-4)(t+2)^2$; valores propios $4, -2$; vectores propios

$$\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \quad \text{para } 4; \quad \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \text{para } -2.$$

   Múltiplos escalares no nulos de la primera; también son posibles las combinaciones lineales del par de vectores propios para -2, o en general, el espacio de soluciones de la ecuación

$$x - y + z = 0.$$

   (c) $(t-2)^2(t-6)$; valores propios 2, 6; vectores propios

$$\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \text{para } 2; \quad \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \quad \text{para } 6.$$

   Pueden ser las combinaciones lineales de los primeros dos. Pueden ser múltiplos escalares no nulos del segundo.

(d) $(t-1)(t-3)^2$; valores propios 1, 3; vectores propios

$$\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \quad \text{para 1;} \quad \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{para 3.}$$

Pueden ser múltiplos escalares del primero. Pueden ser las combinaciones lineales de los dos segundos.

5. (a) Valor propio igual a 1; los vectores propios son múltiplos escalares de ${}^t(1,1)$.
(b) Valor propio igual a 1; los vectores propios son múltiplos escalares de ${}^t(1,0)$.
(c) Valor propio igual a 1; los vectores propios son múltiplos escalares de ${}^t(0,2)$.
(d) Los valores propios son $\lambda_1 = (1+\sqrt{-3})/2$, $\lambda_2 = (1-\sqrt{-3})/2$; los valores propios son múltiplos escalares de ${}^t\big(1,(\lambda+1)^{-1}\big)$, donde $\lambda = \lambda_1$, o $\lambda = \lambda_2$. No hay ningún vector propio real.

6. En todos los casos los valores propios son iguales a 1. Los vectores propios son múltiplos de ${}^t(1,0,0)$.

7. (a) Los valores propios son iguales a $\pm 1, \pm\sqrt{-1}$. Los vectores propios son múltiplos escalares de ${}^t(1,\lambda,\lambda^2,\lambda^3)$, donde λ es cualquiera de los cuatro valores propios. Sólo hay dos valores propios reales.
(b) Los valores propios son 2, $(-1+\sqrt{-3})/2, (-1-\sqrt{-3})/2$. Los vectores propios son múltiplos escalares de ${}^t\left(1, \lambda+1, \dfrac{(\lambda+1)^2+4}{13}\right)$, donde λ es uno de los tres valores propios. Sólo hay un vector propio real, a saber, ${}^t(1,1,3)$, excepto múltiplos escalares reales.

VIII, §3, pág. 233

1. (a) 1, 3
(b) $(1+\sqrt{5})/2, (1-\sqrt{5})/2$
2. (a) 0, 1, 3
(b) $2, 2\pm\sqrt{2}$
El valor máximo de f es el número mayor en cada caso.

3. $\dfrac{-1\pm\sqrt{74}}{2}$

VIII, §4, pág. 236

1. $\sum a_i x_i^2$ si $a_1, \ldots, a_n$ son los elementos de la diagonal.

2. Suponga que B tiene como elementos diagonales a $\lambda_1^{1/2}, \ldots, \lambda_n^{1/2}$.

3. Sea v un vector propio $\neq O$ cuyo valor propio es λ. Entonces $\langle Av, v\rangle = \langle \lambda v, v\rangle = \langle v, v\rangle$. Como $\langle v, v\rangle > 0$ y $\langle Av, v\rangle > 0$, se infiere que $\lambda > 0$. Escoja una base de V que conste de los vectores propios. Entonces el espacio vectorial V se puede identificar como el espacio de vectores de coordenadas con respecto a esta base. La matriz de A es, pues, una matriz diagonal, cuyos elementos diagonales son los valores propios y, por consiguiente, son positivos. Entonces podemos usar el Ejercicio 2 para hallar una raíz cuadrada.

4. Es semejante al Ejercicio 3.

5. A partir de ${}^t(AA) = {}^tA{}^tA = AA$, se infiere que A^2 es simétrica. Además, para $v \neq O$,

$$\langle A^2 v, v\rangle = \langle Av, {}^tAv\rangle = \langle Av, Av\rangle > 0,$$

debido a que $Av \neq O$, puesto que $\langle Av, Av\rangle > O$.

Como ${}^tA^{-1} = A^{-1}$, se infiere que A^{-1} es simétrica. Como A es invertible, un v dado se puede escribir como $v = Aw$ para algún w (a saber, $w = A^{-1}v$). Entonces

$$\langle A^{-1}v, v\rangle = \langle A^{-1}Aw, Aw\rangle = \langle w, Aw\rangle = \langle Aw, w\rangle > 0.$$

En consecuencia, A^{-1} es definitivamente positiva.

6. Suponga (i). A partir de la identidad que aparece en la sugerencia, obtenemos

$$\begin{aligned} 4\langle Uv, Uw\rangle &= \langle U(v+w), U(v+w)\rangle - \langle U(v-w), U(v-w)\rangle \\ &= \langle (v+w), v+w\rangle - \langle (v-w), (v-w)\rangle \\ &= 4\langle v, w\rangle. \end{aligned}$$

Por tanto, $\langle Uv, Uw\rangle = \langle v, w\rangle$. El recíproco es inmediato.

Suponga (i). Entonces $\operatorname{Ker} U = O$, debido a que, si $Uv = O$, entonces $\|v\| = 0$, por lo que $v = O$. Pero una aplicación lineal, cuyo núcleo es O, de un espacio vectorial de dimensión finita, en sí mismo, es invertible, de manera que U es invertible. Además, para todos los v y $w \in V$,

$$\langle Uv, Uw\rangle = \langle {}^tUUv, w\rangle \text{ y también es igual a } \langle v, w\rangle, \text{ por hipótesis.}$$

En consecuencia, ${}^tUU = I$, por lo que ${}^tU = U^{-1}$. Recíprocamente, a partir de ${}^tU = U^{-1}$, obtenemos

$$\langle Uv, Uv\rangle = \langle {}^tUUv, v\rangle = \langle v, v\rangle,$$

por lo que U satisface (i).

7. Como ${}^t({}^tAA) = {}^tA{}^{tt}A = {}^tAA$, se infiere que ${}^tAA$ es simétrica. Además, para $v \neq O$, tenemos que

$$\langle {}^tAAv, v\rangle = \langle Av, Av\rangle > 0$$

debido a que A es invertible, $Av \neq O$. En consecuencia, ${}^tAA$ es definitivamente positiva. Sea $U = AB^{-1}$, donde $B^2 = {}^tAA$ y $BA = AB$, de manera que $B^{-1}A = AB^{-1}$. Entonces

$$\begin{aligned} \langle Uv, Uv\rangle &= \langle AB^{-1}v, AB^{-1}v\rangle = \langle B^{-1}Av, B^{-1}Av\rangle \\ &= \langle Av, {}^tB^{-1}B^{-1}Av\rangle \\ &= \langle v, AA^{-2}Av\rangle = \langle v, v\rangle. \end{aligned}$$

Por tanto, U es unitaria.

8. Sea v un valor propio no nulo $\lambda > 0$. Entonces

$$\langle Bv, Bv\rangle = \lambda^2\langle v, v\rangle = \langle v, v\rangle$$

debido a que B es unitaria. En consecuencia, $\lambda^2 = 1$, por lo que $\lambda = \pm 1$ y, como λ es positivo, $\lambda = 1$. Puesto que V tiene una base que consta de vectores propios, se infiere que $B = I$.

Índice

E

F

G

H

I

L

M

N

O

Este libro, versión en español de *Introduction to Linear Algebra, second edition,* fue concebido como un texto breve para un curso semestral de álgebra lineal. Por ello, la lógica del texto es independiente del cálculo excepto por algún ejemplo o ejercicio ocasional, y se podría utilizar en un curso previo.

En el primer capítulo el autor analiza la conexión fundamental entre el álgebra lineal y la intuición geométrica, y se dan ejemplos concretos de nociones que se desarrollan posteriormente. Después, se comienza con ecuaciones lineales, matrices y la eliminación gaussiana, haciendo énfasis en los aspectos computacionales del álgebra lineal. Por último se trabaja con transformaciones lineales y productos escalares, y sus relaciones con las matrices, combinando así el aspecto computacional con el teórico. Además, se tratan los determinantes —de forma breve, omitiendo varias demostraciones—, y se incluye un capítulo sobre valores y vectores propios como introducción al material de uso constante en las matemáticas y sus aplicaciones.

Este libro, junto con los libros de *Cálculo* e *Introducción al análisis matemático,* forman parte de la trilogía de Lang publicada por Addison-Wesley Iberoamericana, colección indispensable en la biblioteca del estudioso de las matemáticas.



ADDISON-WESLEY IBEROAMERICANA

Apartado Postal 22-456. México 14060, D.F., México
Apartado Aéreo 29696. Bogotá, Colombia
Casilla 70060. Santiago 7, Chile
Apartado Postal 29853. Rio Piedras, Puerto Rico 00929
Martin de los Heros 59, Bis, 4. 28008, Madrid, España
7 Jacob Way. Reading, Massachusetts, E.U.A.